PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

Assistant Commissioner for Patents
United States Patent and Trademark Office
Box PCT
Washington, D.C.20231
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 19 April 2000 (19.04.00)	
International application No. PCT/US99/12385	**Applicant's or agent's file reference** PF-0564 PCT
International filing date (day/month/year) 19 July 1999 (19.07.99)	**Priority date** (day/month/year) 20 July 1998 (20.07.98)
Applicant BANDMAN, Olga et al	

1. The designated Office is hereby notified of its election made:

[X] in the demand filed with the International Preliminary Examining Authority on:

02 February 2000 (02.02.00)

[ ] in a notice effecting later election filed with the International Bureau on:

2. The election [X] was

[ ] was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO **34, chemin des Colombettes** **1211 Geneva 20, Switzerland**	Authorized officer Diana Nissen
Facsimile No.: (41-22) 740.14.35	Telephone No.: (41-22) 338.83.38

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference	**FOR FURTHER ACTION**	see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.
PF-0564 PCT		
International application No.	International filing date *(day/month/year)*	(Earliest) Priority Date *(day/month/year)*
PCT/US 99/ 12385	19/07/1999	20/07/1998
Applicant		
INCYTE PHARMACEUTICALS, INC. et al.		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 6 sheets.

[X] It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**

a. With regard to the **language**, the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.

[ ] the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

b. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international search was carried out on the basis of the sequence listing :

[X] contained in the international application in written form.

[X] filed together with the international application in computer readable form.

[ ] furnished subsequently to this Authority in written form.

[ ] furnished subsequently to this Authority in computer readble form.

[ ] the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

[ ] the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

2. [X] **Certain claims were found unsearchable** (See Box I).

3. [X] **Unity of invention is lacking** (see Box II).

4. With regard to the **title,**

[ ] the text is approved as submitted by the applicant.

[X] the text has been established by this Authority to read as follows:

HUMAN CALCIUM REGULATORY PROTEINS

5. With regard to the **abstract,**

[X] the text is approved as submitted by the applicant.

[ ] the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is Figure No. –

[ ] as suggested by the applicant.

[ ] because the applicant failed to suggest a figure.

[ ] because this figure better characterizes the invention.

[ ] None of the figures.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 99/ 12385

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Remark: Although claims 17-18 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [X] Claims Nos.: 15,16,19
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

See additional sheet

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See additional sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [X] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-21 (partially)

Isolation and cloning of a human cDNA coding for a calcium regulatory protein CaREG1, and the recombinant expression of the same.

2. Claims: 1-21 (partially)

Isolation and cloning of a human cDNA coding for a calcium regulatory protein CaREG2, and the recombinant expression of the same.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 15,16,19

Due to the fact that the agonist and antagonist of claims 15 and 16 were not sufficiently diclosed in the description, a reasonable search could not be carried out for the above mentioned claims together with claim 19.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/12385

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/47 C12N5/10 C07K16/18 A61K38/17
C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMST27 'Online! The Institute for Genomic Research Fetal lung II Homo sapiens cDNA, 18 April 1997 (1997-04-18) XP002121205 see Accesion AA359824 & ADAMS MD ET AL: "Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence" NATURE, vol. 377, 28 September 1995 (1995-09-28), pages 3-17, XP002914403 --- -/--	1-9

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
3 November 1999	16/11/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	ALCONADA RODRIG.., A

8

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMEST1 'Online! National Cancer Institute, Cancer Genome Anatomy Project Homo sapiens cDNA clone, 11 September 1997 (1997-09-11) XP002121206 see Accesion AA551623	1-9
X	DATABASE EMEST25 'Online! National Cancer Institute, Cancer Genome Anatomy Project Homo sapiens cDNA clone, 5 July 1997 (1997-07-05) XP002121207 see Accession AA507028	1-9
X	DATABASE EMEST25 'Online! Wash-Merck EST Project Stratagene endothelial cell H.sapiens cDNA , 15 December 1996 (1996-12-15) XP002121208 see Accession Number AA147831	1-9
A	WO 97 34013 A (HUMAN GENOME SCIENCES INC ;NI JIAN (US); GENTZ REINER L (US); YU G) 18 September 1997 (1997-09-18) page 17 -page 49 claims 1-22	1-16, 19-21
A	WO 98 26068 A (INCYTE PHARMA INC ;GOLI SURYA K (US); HILLMAN JENNIFER L (US)) 18 June 1998 (1998-06-18) page 11 -page 38 claims 1-18	1-16,18, 20,21
A	SCHAFER B W ET AL: "The S100 family of EF-hand calcium-binding proteins: functions and pathology" TIBS TRENDS IN BIOCHEMICAL SCIENCES, vol. 21, no. 4, 1 April 1996 (1996-04-01), page 134-140 XP004050923 ISSN: 0968-0004 the whole document	17-19

INTERNATIONAL SEARCH REPORT

national application No
PCT/US 99/ 12385

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Remark: Although claims 17-18 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [X] Claims Nos.: 15,16,19
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

See additional sheet

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See additional sheet

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [X] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 15,16,19

Due to the fact that the agonist and antagonist of claims 15 and 16 were not sufficiently diclosed in the description, a reasonable search could not be carried out for the above mentioned claims together with claim 19.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-21 (partially)

Isolation and cloning of a human cDNA coding for a calcium regulatory protein CaREG1, and the recombinant expression of the same.

2. Claims: 1-21 (partially)

Isolation and cloning of a human cDNA coding for a calcium regulatory protein CaREG2, and the recombinant expression of the same.

INTE[redacted]TIONAL SEARCH REPORT

...nformation on patent family members

Inte[redacted]onal Application No
PCT/US 99/12385

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9734013 A	18-09-1997	AU 5420796 A	01-10-1997
WO 9826068 A	18-06-1998	US 5763220 A	09-06-1998
		AU 5520398 A	03-07-1998
		EP 0946722 A	06-10-1999
		US 5935931 A	10-08-1999

PATENT COOPERATION TREATY

From the
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To: LUCY J. BILLINGS
INCYTE PHARMACEUTICALS, INC.
3174 PORTER DRIVE
PALO ALTO CA 94304

PCT

NOTIFICATION OF TRANSMITTAL OF INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Rule 71.1)

Date of Mailing *(day/month/year)*	09 NOV 2000
Applicant's or agent's file reference PF-0564 PCT	IMPORTANT NOTIFICATION

International application No.	International filing date *(day/month/year)*	Priority Date *(day/month/year)*
PCT/US99/12385	19 JULY 1999	20 JULY 1998

Applicant
INCYTE PHARMACEUTICALS, INC.

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.

2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.

3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.

4. REMINDER

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices)(Article 39(1))(see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

Name and mailing address of the IPEA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer ANNE L. HOLLRAN
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

Form PCT/IPEA/416 (July 1992)★

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference PF-0564 PCT	**FOR FURTHER ACTION**	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. PCT/US99/12385	International filing date *(day/month/year)* 19 JULY 1999	Priority date *(day/month/year)* 20 JULY 1998
International Patent Classification (IPC) or national classification and IPC Please See Supplemental Sheet.		
Applicant INCYTE PHARMACEUTICALS, INC.		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 9 sheets.

   ☐ This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority. (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of 0 sheets.

3. This report contains indications relating to the following items:

   I ☒ Basis of the report

   II ☐ Priority

   III ☒ Non-establishment of report with regard to novelty, inventive step or industrial applicability

   IV ☒ Lack of unity of invention

   V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

   VI ☐ Certain documents cited

   VII ☐ Certain defects in the international application

   VIII ☐ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
02 FEBRUARY 2000	23 OCTOBER 2000
Name and mailing address of the IPEA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer ANNE L. HOLLERAN Telephone No. (703) 308-0196

Form PCT/IPEA/409 (cover sheet) (July 1998)★

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/US99/12385

I. Basis of the report

1. With regard to the **elements** of the international application:*

[X] the international application as originally filed

[X] the description:
pages 1-51, as originally filed
pages NONE, filed with the demand
pages NONE, filed with the letter of ____

[X] the claims:
pages 52-53, as originally filed
pages NONE, as amended (together with any statement) under Article 19
pages NONE, filed with the demand
pages NONE, filed with the letter of ____

[X] the drawings:
pages 1-9, as originally filed
pages NONE, filed with the demand
pages NONE, filed with the letter of ____

[X] the sequence listing part of the description:
pages 1-4, as originally filed
pages NONE, filed with the demand
pages NONE, filed with the letter of ____

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.
These elements were available or furnished to this Authority in the following language ____ which is:

[ ] the language of a translation furnished for the purposes of international search (under Rule 23.1(b)).

[ ] the language of publication of the international application (under Rule 48.3(b)).

[ ] the language of the translation furnished for the purposes of international preliminary examination (under Rules 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

[X] contained in the international application in printed form.

[X] filed together with the international application in computer readable form.

[ ] furnished subsequently to this Authority in written form.

[ ] furnished subsequently to this Authority in computer readable form.

[ ] The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

[ ] The statement that the information recorded in computer readable form is identical to the writen sequence listing has been furnished.

4. [X] The amendments have resulted in the cancellation of:

[X] the description, pages NONE

[X] the claims, Nos. NONE

[X] the drawings, sheets/~~fig~~ NONE

5. [X] This report has been drawn as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).**

** Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to this report since they do not contain amendments (Rules 70.16 and 70.17).*

***Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.*

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/US99/12385

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

1. The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non obvious), or to be industrially applicable have not been and will not be examined in respect of:

☐ the entire international application.

☒ claims Nos. 15, 16 and 19

because:

☐ the said international application, or the said claim Nos. _ relate to the following subject matter which does not require international preliminary examination *(specify)*.

☐ the description, claims or drawings *(indicate particular elements below)* or said claims Nos. _ are so unclear that no meaningful opinion could be formed *(specify)*.

☐ the claims, or said claims Nos. _ are so inadequately supported by the description that no meaningful opinion could be formed.

☒ no international search report has been established for said claims Nos. 15, 16 and 19 .

2. A meaningful international preliminary examination cannot be carried out due to the failure of the nucleotide and/or amino acid sequence listing to comply with the standard provided for in Annex C of the Administrative Instructions:

☐ the written form has not been furnished or does not comply with the standard.

☐ the computer readable form has not been furnished or does not comply with the standard.

IV. Lack of unity of invention

1. In response to the invitation to restrict or pay additional fees the applicant has:

   ☐ restricted the claims.

   ☐ paid additional fees.

   ☐ paid additional fees under protest.

   ☐ neither restricted nor paid additional fees.

2. ☒ This Authority found that the requirement of unity of invention is not complied with and chose, according to Rule 68.1, not to invite the applicant to restrict or pay additional fees.

3. This Authority considers that the requirement of unity of invention in accordance with Rules 13.1, 13.2 and 13.3 is

   ☐ complied with.

   ☒ not complied with for the following reasons:

   Please See Supplemental Sheet.

4. Consequently, the following parts of the international application were the subject of international preliminary examination in establishing this report:

   ☐ all parts.

   ☒ the parts relating to claims Nos. 1-14, 17-18 and 20-21.

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.
PCT/US99/12385

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. statement

Novelty (N)	Claims	1, 2, 13, 14, 17, 18, 20, 21	YES
	Claims	3-12	NO
Inventive Step (IS)	Claims	1, 2, 13, 14, 17, 18, 20, 21	YES
	Claims	3-12	NO
Industrial Applicability (IA)	Claims	1-14, 17-18, 20-21	YES
	Claims	NONE	NO

2. citations and explanations (Rule 70.7)

Claims 3-12 lack novelty under PCT Article 33(2) as being anticipated by either Accession No. AA551623 (DATABASE EMEST1), Accession No. AA507028 (DATABASE EMEST25), AA147831 (DATABASE EMEST25) or AA359824 (Adams et al).

Claims 3-12 are drawn to polynucleotides which encode polypeptides comprising either SEQ ID NO: 1, SEQ ID NO: 2 or fragments of either sequence. Since "comprising" is understood to have the same meaning as "including", claims 3-7 are interpreted to be the same as polynucleotides encoding polypeptides comprising any size fragment of either SEQ ID NO: 1 or SEQ ID NO: 2 or the same as polynucleotides comprising any size fragment of SEQ ID NO: 3 or SEQ ID NO: 4. Claim 8 is also interpreted as comprising any size fragment since the region over which the percent identity is to be assessed is not specified. Claim 9 which is drawn to complements of the polynucleotides of claim 7 are obvious in view of a disclosure of a sequence that satisfies the limitations of claim 7. The expression vectors, host cells and methods of producing a polypeptide are also obvious over the disclosure of a polynucleotide sequence that satisfies the limitations of claim 3.

Accession No. AA551623 (DATABASE EMEST1), Accession No. AA507028 (DATABASE EMEST25), AA147831 (DATABASE EMEST25) or AA359824 (Adams et al). disclose sequences comprising fragments of polynucleotides encoding either SEQ ID NO: 1 or 2 and therefore provides teachings of nucleotides, vectors, host cells and recombinant methods of making a protein which are the same as that claimed.

Claims 1, 2, and 13-14, 17-18 and 20-21 meet the criteria set out in PCT Article 33(2)-(3), because the prior art does not teach or fairly suggest the polypeptides and pharmaceutical compositions, to the extent that the claims are drawn to polypeptides comprising the entire sequence of either SEQ ID NO: 1 or 2.

Claims 1-14, 17-18 and 20-21 meet the criteria set out in PCT Article 33(4) for industrial applicability.
(Continued on Supplemental Sheet.)

Supplemental Box
(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: Boxes I - VIII

CLASSIFICATION:

The International Patent Classification (IPC) and/or the National classification are as listed below:
IPC(7): C12N 15/12, 5/10; C07K 14/47, 16/18; C12Q 1/68; A61K 38/17 and US Cl.: 530/350, 387.1; 536/23.5; 435/6, 69.1, 325.1, 320.1; 514/2

I. BASIS OF REPORT:

5. (Some) amendments are considered to go beyond the disclosure as filed:
NONE

IV. LACK OF UNITY OF INVENTION:

3. This Authority considers that the requirement of unity of invention in accordance with Rules 13.1, 13.2, and 13.3 is not complied with for the following reasons:
This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1.

Group I, claim(s)1, 2, 13, 17, 18 drawn to polypeptides, pharmaceutical compositions comprising the polypeptides and methods of using the pharmaceutical compositions.
Group II, claim(s) 3-12, drawn to polynucleotides, expression vectors comprising the polynucleotides, host cells and a method of use of the polynucleotides, recombinant production of a polypeptide.
Group III, claim(s) 14, drawn to antibodies which specifically bind to a polypeptide of SEQ ID NO: 1 or SEQ ID NO: 2 or polynucleotides comprising fragments thereof.
Group IV, claims 20 and 21, drawn to methods of detection of a polynucleotide.
The inventions listed as Groups I, II and III do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: each of the groups contains claims to separate products which may be used in different methods and are made by different methods. Group IV contains a second method of use of the the products of group II.

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. SEQ ID NO: 1 and SEQ ID NO: 2 are each separate products. Therefore, the polynucleotides encoding and the antibodies which bind are each separate products.

V. 2. REASONED STATEMENTS - CITATIONS AND EXPLANATIONS (Continued):

---------- NEW CITATIONS ----------
NONE

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12P 21/02, 21/06, C12N 1/20, 15/00, C12Q 1/68, G01N 33/53, C07K 1/00, 14/52, A61K 45/05, 38/00, 38/19

A1

(11) International Publication Number: WO 97/34013

(43) International Publication Date: 18 September 1997 (18.09.97)

(21) International Application Number: PCT/US96/03219

(22) International Filing Date: 11 March 1996 (11.03.96)

(71) Applicant *(for all designated States except US)*: HUMAN GENOME SCIENCES, INC. [US/US]; 9410 Key West Avenue, Rockville, MD 20850-3338 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: NI, Jian [CN/US]; 5502 Manorfield Road, Rockville, MD 20853 (US). YU, Guo-Liang [CN/US]; 13524 Straw Bale Lane, Damestown, MD 20878 (US). GENTZ, Reiner, L. [DE/US]; 13404 Fairland Park Drive, Silver Spring, MD 20904 (US).

(74) Agents: BENSON, Robert, H. et al.; Human Genome Sciences, Inc., 9410 Key West Avenue, Rockville, MD 20850 (US).

(81) Designated States: AM, AT, AU, BG, BR, BY, CA, CN, CZ, EE, FI, GE, HU, JP, KG, KP, KR, KZ, LT, LV, MD, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, TJ, TM, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: CHEMOTACTIC CYTOKINE II

(57) Abstract

The invention relates to CCII polypeptides, polynucleotides encoding the polypeptides, methods for producing the polypeptides, in particular by expressing the polynucleotides, and agonists and antagonists of the polypeptides. The invention further relates to methods for utilizing such polynucleotides, polypeptides, agonists and antagonists for applications, which relate, in part, to research, diagnostic and clinical arts.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

CHEMOTACTIC CYTOKINE II

This invention relates, in part, to newly identified polynucleotides and polypeptides; variants and derivatives of the polynucleotides and polypeptides; processes for making the polynucleotides and the polypeptides, and their variants and derivatives; agonists and antagonists of the polypeptides; and uses of the polynucleotides, polypeptides, variants, derivatives, agonists and antagonists. In particular, in these and in other regards, the invention relates to polynucleotides and polypeptides of human Chemotactic cytokine II, sometimes hereinafter referred to as "CCII".

BACKGROUND OF THE INVENTION

The cytokine family of proteins exhibits a wide variety of functions. A hallmark feature is their ability to elicit chemotactic migration of distinct cell types, including polymorphonuclear cells and macrophages. Many cytokines have proinflammatory activity and are involved in multiple steps during inflammatory reactions. In addition to their involvement in inflammation, cytokines have been shown to exhibit other activities. For example, interleukin-8 (IL-8) promotes proliferation of keratinocytes.

In light of the diverse biological activities, it is not surprising that cytokines have been implicated in a number of physiological and disease conditions, including cell migration, for example, lymphocyte trafficking, wound healing, hematopoietic

regulation and immunological reactions such as allergy, asthma and arthritis.

The S100 family of calcium binding proteins has chemotactic activity for polymorphonuclear cells, mononuclear cells and neutrophils and are calcium binding proteins. The S100 protein has been recently identified in cells of myeloid origin and consists of macrophage inhibitory factor-related protein (MRP-8), MRP-14, chemotactic protein 10 (CP-10) and calgranulin C.

MRP-8 and MRP-14 are expressed in a cell lineage-specific manner. Alignment of individual sequences shows that there is overall conservation of structure within the family, a notable feature being the two calcium binding sites, which are the "EF hand" type. Sequences at both the NH_2- and COOH- terminal ends of MRP-8 and MRP-14 are relatively hydrophobic. An attractive hypothesis is that these regions of the molecule are buried until calcium binding brings about the conformational changes that cause their exposure, making them potentially available for interactions with other effector molecules. Because of the extended sequence of its COOH- terminal "tail" MRP-14 is the largest member of the S100 family. (Hessian, P., *et al.*, J. Leuk. Bio., 53:197-204 (1993).

Each gene in the S100 family is composed of three exons with one intron interrupting the protein-coding sequence between the two EF hands. The MRP-8 and MRP-14 genes are both localized to chromosome 1Q12-Q21 with an undefined distance between them (Dorin, J.R., *et al.*, Nature, 326:614-617 (1987) and Lagasse, E. and Clerc, R.G., Mol. Cell Biol., 8:2402, 2410 (1988)). Two other S100 family members 1882 (CAPL) and calcyclin/2A9 (CACY) also map to chromosome 1Q12-Q21 (Dorian, J.R., *et al.*, Genomics, 8:420-426 (1990). It is probably that co-segregation of these five genes on chromosome 1 may represent an S100 family locus. However, this does not apply to all S100 homologs.

MRP-8 and MRP-14 are restricted to cells of the monocytes/macrophage lineage, neutrophils, and under certain circumstance keratinocytes, suggesting that its expression is tightly regulated during differentiation (Hogg, N., *et al.*, Eur. J. Immunol., 19:1053-1061 (1989)). Thus, monocytes and neutrophils in the circulation express MRP-8 and MRP-14, in contrast to other

related cells such as lymphocytes, platelets, basophils and eosinophils which do not (Id.).

Resident tissue macrophages do not express MRP-8 and MRP-14, implying that differentiation of monocytes to macrophages is normally associated with loss of this molecule (Id.). Furthermore, immunohistochemical data show that at inflammatory sites MRP-8 and MRP-14 positive cells are associated with vessels but that the majority of monocytes already within the tissues at these sites have lost MRP-8 and MRP-14 expression (Id.). In keeping with these observations, tissue culture-matured monocytes down regulate this molecule (Zwadlo, G., *et al.*, Clin. Exp. Immunol., 72:510-515 (1988)).

At sites of chronic inflammation in patients with diseases such as rheumatoid arthritis, sarcoidosis, tuberculosis or onchocerciasis macrophages express both MRP-8 and MRP-14 (Palmer, D.G., *et al.*, Clin. Immunol. Immunopathol., 45:17-28 (1987)). In contrast, macrophages in acutely inflamed tissues may express only MRP-14 (Delabie, J., *et al.*, Clin. Exp. Immunol., 81:123-126 (1990)). The expression of MRP-8 and MRP-14 by macrophages could be flecked exposure to tissue stimuli that either maintain expression or induce re-expression of the molecule (Palmer, D.G., *et al.*, Clin. Immunol. Immunopathol., 45:17-28 (1987)).

In common with other members of the S100 family, MRP-8 and MRP-14 are found predominately in a cytosolic location in both monocytes and neutrophils (Dale, I., *et al.*, Eur. J. Biochem., 134:1-6 (1983)). It is also possible that MRP-8 and MRP-14 can be expressed on the cell surface, although the majority of antibodies specific for these proteins do not react with circulating monocytes or neutrophils. There is also evidence that MRP-8 and MRP-14 exist extracellularly, however, neither protein has the signal peptide sequence for membrane translocation. Thus, MRP-8 and MRP-14 fall into the category of proteins, including interleukin-1 and basic fibroblast growth factor, that clearly have extracellular functions but about which little is known of their cellular release. Finally, MRP-8 and MRP-14 are found in the serum of patients with cystic fibrosis and other chronic inflammatory states such as rheumatoid arthritis and sarcoidosis (Bullock, S., *et al.*, Clin. Genet., 21:336-341 (1982)). MRP-14 also belongs to a novel subfamily of highly homologous calcium-binding proteins which

includes S100 alpha, S100 beta, intestinal calcium-binding protein, P11 and calcyclin (2A9).

CP-10 is one of the most potent chemotactic cytokines of the S100 family. The CCII gene of the present invention has homology to murine CP-10, purified from supernatants of activated murine spleen cells (Lackman, M., *et al.*, J. Biol. Chem., 267:7499 (1992)). An extracellular function of the murine CP-10 includes a potent chemotactic agent for murine and human polymorphonuclear cells (PMN) and murine monocytes and is involved in phagocyte recruitment during inflammatory reactions. CP-10 has maximal chemotactic activity for neutrophils at 10(-13) M. The 76 amino acid sequence revealed up to 55% sequence homology with S100, Ca^{++}-binding proteins. A combination of Western and Northern analyses identified CP-10 in murine peritoneal exudate PMN and macrophages, splenocytes, bone marrow cells, and WEHI-265 cells (all of myeloid origin), but not in thymus, liver, lung, 3T3 fibroblasts, EL4 lymphoma cells, or bEND 3 brain endothelial cells, indicating cell-specific regulation of CP-10 expression. CP-10 has an apparent molecular weight of 10.3 kd and a complete sequence of 88 amino acids.

S100 proteins are characterized by two calcium binding regions, which are strongly conserved and are separated by an 8 to 12 amino acid hinge region (Kligman, D., Trends Biochem. Sci., 13:437 (1988)). Although the hinge region length is conserved, the amino acid sequences are widely divergent. This divergence led to the hypothesis that the hinge region may concur functional specificity by interaction with the factor proteins (Id.).

SUMMARY OF THE INVENTION

Toward these ends, and others, it is an object of the present invention to provide polypeptides, inter alia, that have been identified as novel cytokines by homology between the amino acid sequence set out in Figure 1 (SEQ ID NO:2) and known amino acid sequences of other proteins such as CP-10 proteins.

It is a further object of the invention, moreover, to provide polynucleotides that encode CCII, particularly polynucleotides that encode the polypeptide herein designated CCII.

In a particularly preferred embodiment of this aspect of the invention the polynucleotide comprises the region encoding human CCII in the sequence set out in Figure 1 (SEQ ID NO:1).

In accordance with this aspect of the present invention there is provided an isolated nucleic acid molecule encoding a mature polypeptide expressed by the human cDNA contained in ATCC Deposit No. 97405.

In accordance with this aspect of the invention there are provided isolated nucleic acid molecules encoding human CCII, including mRNAs, cDNAs, genomic DNAs and, in further embodiments of this aspect of the invention, biologically, diagnostically, clinically or therapeutically useful variants, analogs or derivatives thereof, or fragments thereof, including fragments of the variants, analogs and derivatives.

Among the particularly preferred embodiments of this aspect of the invention are naturally occurring allelic variants of human CCII.

It also is an object of the invention to provide CCII polypeptides, particularly human CCII polypeptides, that may be employed to treat tumors, chronic infections, leukemia, T-cell mediated auto-immune diseases, parasitic infections, psoriasis, asthma, allergy, to regulate hematopoiesis, to stimulate growth factor activity, to inhibit angiogenesis, to promote wound healing, to treat inflammatory disorders, to control cellular immune reactions, to treat malignant diseases, and to inhibit casein kinase to activity.

In accordance with this aspect of the invention there are provided novel polypeptides of human origin referred to herein as CCII as well as biologically, diagnostically or therapeutically useful fragments, variants and derivatives thereof, variants and derivatives of the fragments, and analogs of the foregoing.

Among the particularly preferred embodiments of this aspect of the invention are variants of human CCII encoded by naturally occurring alleles of the human CCII gene.

It is another object of the invention to provide a process for producing the aforementioned polypeptides, polypeptide fragments, variants and derivatives, fragments of the variants and derivatives, and analogs of the foregoing. In a preferred embodiment of this aspect of the invention there are provided

methods for producing the aforementioned CCII polypeptides comprising culturing host cells having expressibly incorporated therein an exogenously-derived human CCII-encoding polynuclectide under conditions for expression of human CCII in the host and then recovering the expressed polypeptide.

In accordance with another object the invention there are provided products, compositions, processes and methods that utilize the aforementioned polypeptides and polynucleotides for research, biological, clinical and therapeutic purposes, *inter alia*.

In accordance with certain preferred embodiments of this aspect of the invention, there are provided products, compositions and methods, *inter alia*, for, among other things: assessing CCII expression in cells by determining CCII polypeptides or CCII-encoding mRNA; stimulating migration of PMN's and stimulating cellular immune reactions, in vitro, ex vivo or in vivo by exposing cells to CCII polypeptides or polynucleotides as disclosed herein; assaying genetic variation and aberrations, such as defects, in CCII genes; and administering a CCII polypeptide or polynucleotide to an organism to augment CCII function or remediate CCII dysfunction.

In accordance with certain preferred embodiments of this and other aspects of the invention there are provided probes that hybridize to human CCII sequences.

In certain additional preferred embodiments of this aspect of the invention there are provided antibodies against CCII polypeptides. In certain particularly preferred embodiments in this regard, the antibodies are highly selective for human CCII.

In accordance with another aspect of the present invention, there are provided CCII agonists. Among preferred agonists are molecules that mimic CCII, that bind to CCII-binding molecules or receptor molecules, and that elicit or augment CCII-induced responses. Also among preferred agonists are molecules that interact with CCII or CCII polypeptides, or with other modulators of CCII activities, and thereby potentiate or augment an effect of CCII or more than one effect of CCII.

In accordance with yet another aspect of the present invention, there are provided CCII antagonists. Among preferred antagonists are those which mimic CCII so as to bind to CCII

receptor or binding molecules but not elicit a CCII-induced response or more than one CCII-induced response or which prevent expression of CCII. Also among preferred antagonists are molecules that bind to or interact with CCII so as to inhibit an effect of CCII or more than one effect of CCII. The agonists and antagonists may be used to mimic, augment or inhibit the action of CCII polypeptides.

Antagonists may be employed to treat certain auto-immune diseases, atherosclerosis, chronic inflammatory and infectious diseases, histamine and IgE-mediated allergic reactions, prostaglandin-independent fever, bone marrow failure, cancers, silicosis, sarcoidosis, rheumatoid arthritis, shock, hyper-eosinophilic syndrome and fibrosis in the asthmatic lung, cystic fibrosis, malignant diseases, psoriasis, diapedesis and urinary and kidney stones.

In a further aspect of the invention there are provided compositions comprising a CCII polynucleotide or a CCII polypeptide for administration to cells in vitro, to cells ex vivo and to cells in vivo, or to a multicellular organism. In certain particularly preferred embodiments of this aspect of the invention, the compositions comprise a CCII polynucleotide for expression of a CCII polypeptide in a host organism for treatment of disease. Particularly preferred in this regard is expression in a human patient for treatment of a dysfunction associated with aberrant endogenous activity of CCII.

Other objects, features, advantages and aspects of the present invention will become apparent to those of skill from the following description. It should be understood, however, that the following description and the specific examples, while indicating preferred embodiments of the invention, are given by way of illustration only. Various changes and modifications within the spirit and scope of the disclosed invention will become readily apparent to those skilled in the art from reading the following description and from reading the other parts of the present disclosure.

BRIEF DESCRIPTION OF THE DRAWINGS

The following drawings depict certain embodiments of the invention. They are illustrative only and do not limit the invention otherwise disclosed herein.

Figure 1 shows the nucleotide and deduced amino acid sequence of human CCII.

Figure 2 shows the regions of similarity between amino acid sequences of CCII and murine CP-10 polypeptide (SEQ ID NO:9).

Figure 3 shows structural and functional features of CCII deduced by the indicated techniques, as a function of amino acid sequence.

GLOSSARY

The following illustrative explanations are provided to facilitate understanding of certain terms used frequently herein, particularly in the examples. The explanations are provided as a convenience and are not limitative of the invention.

DIGESTION of DNA refers to catalytic cleavage of the DNA with a restriction enzyme that acts only at certain sequences in the DNA. The various restriction enzymes referred to herein are commercially available and their reaction conditions, cofactors and other requirements for use are known and routine to the skilled artisan.

For analytical purposes, typically, 1 μg of plasmid or DNA fragment is digested with about 2 units of enzyme in about 20 μl of reaction buffer. For the purpose of isolating DNA fragments for plasmid construction, typically 5 to 50 μg of DNA are digested with 20 to 250 units of enzyme in proportionately larger volumes.

Appropriate buffers and substrate amounts for particular restriction enzymes are described in standard laboratory manuals, such as those referenced below, and they are specified by commercial suppliers.

Incubation times of about 1 hour at 37°C are ordinarily used, but conditions may vary in accordance with standard procedures, the supplier's instructions and the particulars of the reaction. After digestion, reactions may be analyzed, and fragments may be purified by electrophoresis through an agarose or polyacrylamide gel, using well known methods that are routine for those skilled in the art.

GENETIC ELEMENT generally means a polynucleotide comprising a region that encodes a polypeptide or a region that regulates transcription or translation or other processes important to expression of the polypeptide in a host cell, or a polynucleotide

comprising both a region that encodes a polypeptide and a region operably linked thereto that regulates expression.

Genetic elements may be comprised within a vector that replicates as an episomal element; that is, as a molecule physically independent of the host cell genome. They may be comprised within mini-chromosomes, such as those that arise during amplification of transfected DNA by methotrexate selection in eukaryotic cells. Genetic elements also may be comprised within a host cell genome; not in their natural state but, rather, following manipulation such as isolation, cloning and introduction into a host cell in the form of purified DNA or in a vector, among others.

ISOLATED means altered [from] "by the hand of man" from its natural state; i.e., that, if it occurs in nature, it has been changed or removed from its original environment, or both.

For example, a naturally occurring polynucleotide or a polypeptide naturally present in a living animal in its natural state is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. For example, with respect to polynucleotides, the term isolated means that it is separated from the chromosome and cell in which it naturally occurs.

As part of or following isolation, such polynucleotides can be joined to other polynucleotides, such as DNAs, for mutagenesis, to form fusion proteins, and for propagation or expression in a host, for instance. The isolated polynucleotides, alone or joined to other polynucleotides such as vectors, can be introduced into host cells, in culture or in whole organisms. Introduced into host cells in culture or in whole organisms, such DNAs still would be isolated, as the term is used herein, because they would not be in their naturally occurring form or environment. Similarly, the polynucleotides and polypeptides may occur in a composition, such as a media formulations, solutions for introduction of polynucleotides or polypeptides, for example, into cells, compositions or solutions for chemical or enzymatic reactions, for instance, which are not naturally occurring compositions, and,

therein remain isolated polynucleotides or polypeptides within the meaning of that term as it is employed herein.

LIGATION refers to the process of forming phosphodiester bonds between two or more polynucleotides, which most often are double stranded DNAs. Techniques for ligation are well known to the art and protocols for ligation are described in standard laboratory manuals and references, such as, for instance, Sambrook et al., MOLECULAR CLONING, A LABORATORY MANUAL, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1989) and Maniatis et al., pg. 146, as cited below.

OLIGONUCLEOTIDE(S) refers to relatively short polynucleotides. Often the term refers to single-stranded deoxyribonucleotides, but it can refer as well to single-or double-stranded ribonucleotides, RNA:DNA hybrids and double-stranded DNAs, among others.

Oligonucleotides, such as single-stranded DNA probe oligonucleotides, often are synthesized by chemical methods, such as those implemented on automated oligonucleotide synthesizers. However, oligonucleotides can be made by a variety of other methods, including in vitro recombinant DNA-mediated techniques and by expression of DNAs in cells and organisms.

Initially, chemically synthesized DNAs typically are obtained without a 5′ phosphate. The 5′ ends of such oligonucleotides are not substrates for phosphodiester bond formation by ligation reactions that employ DNA ligases typically used to form recombinant DNA molecules. Where ligation of such oligonucleotides is desired, a phosphate can be added by standard techniques, such as those that employ a kinase and ATP.

The 3′ end of a chemically synthesized oligonucleotide generally has a free hydroxyl group and, in the presence of a ligase, such as T4 DNA ligase, readily will form a phosphodiester bond with a 5′ phosphate of another polynucleotide, such as another oligonucleotide. As is well known, this reaction can be prevented selectively, where desired, by removing the 5′ phosphates of the other polynucleotide(s) prior to ligation.

PLASMIDS generally are designated herein by a lower case p preceded and/or followed by capital letters and/or numbers, in accordance with standard naming conventions that are familiar to those of skill in the art.

Starting plasmids disclosed herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids by routine application of well known, published procedures. Many plasmids and other cloning and expression vectors that can be used in accordance with the present invention are well known and readily available to those of skill in the art. Moreover, those of skill readily may construct any number of other plasmids suitable for use in the invention. The properties, construction and use of such plasmids, as well as other vectors, in the present invention will be readily apparent to those of skill from the present disclosure.

POLYNUCLEOTIDE(S) generally refers to any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. Thus, for instance, polynucleotides as used herein refers to, among others, single-and double-stranded DNA, DNA that is a mixture of single-and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, polynucleotide as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide.

As used herein, the term polynucleotide includes DNAs or RNAs as described above that contain one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "polynucleotides" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are polynucleotides as the term is used herein.

It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term polynucleotide as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of polynucleotides, as well as the

chemical forms of DNA and RNA characteristic of viruses and cells, including simple and complex cells, inter alia.

POLYPEPTIDES, as used herein, includes all polypeptides as described below. The basic structure of polypeptides is well known and has been described in innumerable textbooks and other publications in the art. In this context, the term is used herein to refer to any peptide or protein comprising two or more amino acids joined to each other in a linear chain by peptide bonds. As used herein, the term refers to both short chains, which also commonly are referred to in the art as peptides, oligopeptides and oligomers, for example, and to longer chains, which generally are referred to in the art as proteins, of which there are many types.

It will be appreciated that polypeptides, as is well known and as the term is used herein, generally are formed of the 20 naturally occurring amino acids, and that the amino acids in a polypeptide generally are joined to one another in a linear chain by peptide bonds between the alpha carboxyl and the alpha amino groups of adjacent, succeeding amino acids.

By convention, the sequence of amino acids in a chain usually, but not always, is written beginning (on the left and at the top) with the amino acid having a free alpha amino group. This amino acid is taken as the amino terminus of the polypeptide, also referred to as the N-terminus. Each successive amino acid then is listed in turn, ending with the amino acid having a free carboxyl group (at bottom and right), which is taken as the carboxyl terminus of the polypeptide, also called the C-terminus.

Individual amino acids in a polypeptide commonly are referred to as amino acid residues, and as residues. Generally, the amino acids in a polypeptide are numbered beginning with the amino terminus and proceeding integer by integer and residue by residue to the carboxyl terminus. However, for polypeptides that first are synthesized in cells as precursors to a mature form, it also is common to begin numbering amino acids with the first residue of the mature form. Then, the upstream residues (i.e., those closer to the N-terminus) are assigned negative numbers counting back from residue one (the N-terminus of the mature form) to the N-terminus of the earliest precursor form. Other numbering schemes also have been employed, but less commonly.

Notwithstanding the foregoing general characteristics, it will be appreciated that polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally occurring amino acids, and that many amino acids, including the terminal amino acids, may be modified in a given polypeptide, either by natural processes, such as processing and other post-translational modifications, but also by chemical modification techniques which are well known to the art. Even the common modifications that occur naturally in polypeptides are too numerous to list exhaustively here, but they are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature, and they are well known to those of skill in the art. Among the known modifications which may be present in polypeptides of the present are, to name an illustrative few, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Such modifications are well known to those of skill and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such as, for instance PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as, for example, those provided by Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York (1983);

Seifter et al., Analysis for protein modifications and nonprotein cofactors, Meth. Enzymol. 182: 626-646 (1990) and Rattan et al., Protein Synthesis: Posttranslational Modifications and Aging, Ann. N.Y. Acad. Sci. 663: 48-62 (1992).

It will be appreciated, as is well known and as noted above, that polypeptides are not always entirely linear. For instance, polypeptides may be branched as a result of ubiquitination, and they may be circular, with or without branching, generally as a result of posttranslation events, including natural processing event and events brought about by human manipulation which do not occur naturally. Circular, branched and branched circular polypeptides may be synthesized by non-translation natural process and by entirely synthetic methods, as well.

Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. In fact, blockage of the amino or carboxyl group in a polypeptide, or both, by a covalent modification, is common in naturally occurring and synthetic polypeptides and such modifications may be present in polypeptides of the present invention, as well. For instance, the amino terminal residue of polypeptides made in E. coli, prior to proteolytic processing, almost invariably will be N-formylmethionine.

The modifications that occur in a polypeptide often will be a function of how it is made. For polypeptides made by expressing a cloned gene in a host, for instance, the nature and extent of the modifications in large part will be determined by the host cell posttranslational modification capacity and the modification signals present in the polypeptide amino acid sequence. For instance, as is well known, glycosylation often does not occur in bacterial hosts such as E. coli. Accordingly, when glycosylation is desired, a polypeptide should be expressed in a glycosylating host, generally a eukaryotic cell. Insect cell often carry out the same posttranslational glycosylations as mammalian cells and, for this reason, insect cell expression systems have been developed to express efficiently mammalian proteins having native patterns of glycosylation, inter alia. Similar considerations apply to other modifications.

It will be appreciated that the same type of modification may be present in the same or varying degree at several sites in a

given polypeptide. Also, a given polypeptide may contain many types of modifications.

In general, as used herein, the term polypeptide encompasses all such modifications, particularly those that are present in polypeptides synthesized by expressing a polynucleotide in a host cell.

VARIANT(S) of polynucleotides or polypeptides, as the term is used herein, are polynucleotides or polypeptides that differ from a reference polynucleotide or polypeptide, respectively. Variants in this sense are described below and elsewhere in the present disclosure in greater detail.

(1) A polynucleotide that differs in nucleotide sequence from another, reference polynucleotide. Generally, differences are limited so that the nucleotide sequences of the reference and the variant are closely similar overall and, in many regions, identical.

As noted below, changes in the nucleotide sequence of the variant may be silent. That is, they may not alter the amino acids encoded by the polynucleotide. Where alterations are limited to silent changes of this type a variant will encode a polypeptide with the same amino acid sequence as the reference. Also as noted below, changes in the nucleotide sequence of the variant may alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Such nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below.

(2) A polypeptide that differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference and the variant are closely similar overall and, in many region, identical.

A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions, fusions and truncations, which may be present in any combination.

Among preferred variants are those that vary from a reference by conservative amino acid substitutions. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Typically seen as

conservative substitutions are the replacements, one for another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr.

RECEPTOR MOLECULE, as used herein, refers to molecules which bind or interact specifically with CCII polypeptides of the present invention, including not only classic receptors, which are preferred, but also other molecules that specifically bind to or interact with polypeptides of the invention (which also may be referred to as "binding molecules" and "interaction molecules," respectively and as "CCII binding molecules" and "CCII interaction molecules." Binding between polypeptides of the invention and such molecules, including receptor or binding or interaction molecules may be exclusive to polypeptides of the invention, which is very highly preferred, or it may be highly specific for polypeptides of the invention, which is highly preferred, or it may be highly specific to a group of proteins that includes polypeptides of the invention, which is preferred, or it may be specific to several groups of proteins at least one of which includes polypeptides of the invention.

Such molecules generally are proteins, which may be single or multichain proteins and multisubunit or multiprotein complexes, such as those of classic cell surface receptors, which are highly preferred in the invention. Receptor molecules also may be non-protein molecules that bind to or interact specifically with polypeptides of the invention.

Such molecules may occur in membranes, such as classic cell surface receptors, or they may occur intracellularly, in the cytosol, inside organelles, or in the surface of organelles, for instance. Among particularly preferred receptor molecules in this regard are membrane bound receptors, particularly cell membrane receptors, especially cell surface receptors. Also among preferred receptors are those that occur in the membranes of organelles, particularly nuclear membrane receptors and mitochondrial membrane receptors.

Receptors also may be non-naturally occurring, such as antibodies and antibody-derived reagents that bind specifically to polypeptides of the invention.

DESCRIPTION OF THE INVENTION

The present invention relates to novel CCII polypeptides and polynucleotides, among other things, as described in greater detail below. In particular, the invention relates to polypeptides and polynucleotides of a novel human CCII, which is related by amino acid sequence homology to murine CP-10

The invention relates especially to CCII having the nucleotide and amino acid sequences set out in Figure 1 (SEQ ID NOS:1 and 2) , and to the CCII nucleotide and amino acid sequences of the human cDNA in ATCC Deposit No. 97405. It will be appreciated that the nucleotide and amino acid sequences set out in Figure 1 (SEQ ID NOS:1 and 2) were obtained by sequencing the human cDNA of the deposited clone. Hence, the sequence of the deposited clone is controlling as to any discrepancies between the two and any reference to the sequences of Figure 1 (SEQ ID NO:1) include reference to the sequence of the human cDNA of the deposited clone.

Polynucleotides

In accordance with one aspect of the present invention, there are provided isolated polynucleotides which encode the CCII polypeptide having the deduced amino acid sequence of Figure 1 (SEQ ID NO:2).

Using the information provided herein, such as the polynucleotide sequence set out in Figure 1 (SEQ ID NO:1), a polynucleotide of the present invention encoding human CCII polypeptide may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA from cells of a human fetal kidney as starting material.

Human CCII of the invention is structurally related to other proteins of the SP100 family, as shown by the results of sequencing the cDNA encoding human CCII in the deposited clone. The cDNA sequence thus obtained is set out in Figure 1 (SEQ ID NO:1). It contains an open reading frame encoding a protein of about 98 amino acid residues with a deduced molecular weight of about 11471.30 Daltons. The protein exhibits greatest homology to murine CP-10

among known proteins. The amino acid residues of the CCII of Figure 1 (SEQ ID NO:2) have about 20.225 % identity and about 52.809 % similarity with the amino acid sequence of murine CP-10.

Polynucleotides of the present invention may be in the form of RNA, such as mRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced by chemical synthetic techniques or by a combination thereof. The DNA may be double-stranded or single-stranded. Single-stranded DNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

A polynucleotide of the present invention may a naturally occurring sequence, such as that of a naturally occurring allelic variant, or it may have a sequence that does not occur in nature, such as a sequence that has been produced, for instance, by in vitro mutagenesis techniques.

The coding sequence which encodes the polypeptide may be identical to the coding sequence of the polynucleotide shown in Figure 1 (SEQ ID NO:1). It also may be a polynucleotide with a different sequence, which, as a result of the redundancy (degeneracy) of the genetic code, encodes the polypeptide of the DNA of Figure 1 (SEQ ID NO:1).

Polynucleotides of the present invention which encode the polypeptide of Figure 1 (SEQ ID NO:2) may include, but are not limited to the coding sequence for the mature polypeptide, by itself; the coding sequence for the mature polypeptide and additional coding sequences, such as those encoding a leader or secretory sequence, such as a pre-, or pro- or prepro- protein sequence; the coding sequence of the mature polypeptide, with or without the aforementioned additional coding sequences, together with additional, non-coding sequences, including for example, but not limited to introns and non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences that play a role in transcription, mRNA processing - including splicing and polyadenylation signals, for example -ribosome binding and stability of mRNA.

In accordance with the foregoing, the term "polynucleotide encoding a polypeptide" as used herein encompasses polynucleotides which include a sequence encoding a polypeptide of the present

invention, particularly the human CCII having the amino acid sequence set out in Figure 1 (SEQ ID NO:2). The term encompasses polynucleotides that include a single continuous region or discontinuous regions encoding the polypeptide, together with additional regions, that also may contain coding and/or non-coding sequences.

The present invention further relates to variants of the herein above described polynucleotides which encode for fragments, analogs and derivatives of the polypeptide having the deduced amino acid sequence of Figure 1 (SEQ ID NO:2). A variant of the polynucleotide may be a naturally occurring variant such as a naturally occurring allelic variant, or it may be a variant that is not known to occur naturally. Such non-naturally occurring variants of the polynucleotide may be made by mutagenesis techniques, including those applied to polynucleotides, cells or organisms.

The present invention includes polynucleotides encoding the same mature polypeptide as shown in Figure 1 (SEQ ID NO:2). Further, the invention includes variants of such polynucleotides that encode a fragment, derivative or analog of the polypeptide of Figure 1 (SEQ ID NO:2). Among variants in this regard are variants that differ from the aforementioned polynucleotides by nucleotide substitutions, deletions or additions. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding or non-coding regions or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions.

Variants of the invention may have a sequence that occurs in nature or they may have a sequence that does not occur naturally. As herein above indicated, the polynucleotide may have a coding sequence which is a naturally occurring allelic variant of the coding sequence shown in Figure 1 (SEQ ID NO:1). As known in the art, an allelic variant is an alternate form of a polynucleotide sequence which may have a substitution, deletion or addition of one or more nucleotides.

Among the particularly preferred embodiments of the invention in this regard are polynucleotides encoding polypeptides having the amino acid sequence of CCII set out in Figure 1 (SEQ ID NO:2);

variants, analogs, derivatives and fragments thereof, and fragments of the variants, analogs and derivatives.

Further particularly preferred in this regard are polynucleotides encoding CCII variants, analogs, derivatives and fragments, and variants, analogs and derivatives of the fragments, which have the amino acid sequence of the CCII polypeptide of Figure 1 (SEQ ID NO:2) in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, deleted or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the CCII. Also especially preferred in this regard are conservative substitutions. Most highly preferred are polynucleotides encoding polypeptides having the amino acid sequence of Figure 1 (SEQ ID NO:2) without substitutions.

Further preferred embodiments of the invention are polynucleotides that are more than 70% identical to a polynucleotide encoding the CCII polypeptide having the amino acid sequence set out in Figure 1 (SEQ ID NO:2), and polynucleotides which are complementary to such polynucleotides. Alternatively, most highly preferred are polynucleotides that comprise a region that is at least 70% identical to a polynucleotide encoding the CCII polypeptide of the human cDNA of the deposited clone. In this regard, polynucleotides at least 90% identical to the same are particularly preferred, and among these particularly preferred polynucleotides, those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred among those with at least 95%, and among these those with at least 98% and at least 99% are particularly highly preferred, with at least 99% being the more preferred.

Also particularly preferred in this regard are polynucleotides encoding a polypeptide having the amino acid sequence of the CCII set out in Figure 1 (SEQ ID NO:2). As set out elsewhere herein, the polynucleotide may encode the polypeptide in a continuous region or in a plurality of two or more discontinuous exons, and it may comprise additional regions as well, which are unrelated to the coding region or regions.

Most highly preferred in this regard are polynucleotides that comprise a region that are at least 70% identical to the CCII-

encoding portion of the polynucleotide set out in Figure 1 (SEQ ID NO:1). Alternatively, most highly preferred are polynucleotides that comprise a region that is at least 85% identical to the CCII-encoding portion of the human cDNA the deposited clone. Among such polynucleotides, those at least 90% identical to the same are particularly preferred, and, among these particularly preferred polynucleotides, those with at least 95% are especially preferred. Furthermore, those with at least 97% are highly preferred among those with at least 95% and among these those with at least 98% and at least 99% are particularly highly preferred, with at least 99% being the more preferred of these.

The present invention also includes polynucleotides in which the sequence encoding the mature polypeptide is fused in the same reading frame to additional sequences. Such sequences include signal sequences, which facilitate transport of the nascent protein into the endoplasmic reticulum, pro-sequences that are associated with inactive precursor forms of the polypeptide, which may facilitate trafficking of the protein in a cell or out of a cell or may improve persistence of the protein in a cell or in an extracellular compartment. Such sequences also may be added to facilitate production and purification, or to add additional functional domains, as discussed elsewhere herein. Thus, polynucleotides of the invention may encode, in addition to a mature cytokine, particularly CCII, for example, a leader sequence, such as a signal peptide which functions as a secretory sequence for controlling transport of the polypeptide into the lumen of the endoplasmic reticulum. The leader sequence may be removed by the host cell, as is generally the case for signal peptides, yielding another precursor protein or the mature polypeptide. A precursor protein having a leader sequence often is called a preprotein.

The polynucleotides also may encode a polypeptide which is the mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids interior to the mature polypeptide (when the mature form has more than one polypeptide chain, for instance). Such sequences may play a role in processing of a protein from precursor to a mature form, may facilitate protein trafficking, may prolong or shorten protein half-life or may facilitate manipulation of a protein for assay or production, among other things. As

generally is the case in situ, the additional amino acids may be processed away from the mature protein by cellular enzymes.

A precursor protein, having the mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. When prosequences are removed such inactive precursors generally are activated. Some or all of the prosequences may be removed before activation. Generally, such precursors are called proproteins.

In sum, a polynucleotide of the present invention may encode a mature protein, a mature protein plus a leader sequence (which may be referred to as a preprotein), a precursor of a mature protein having one or more prosequences which are not the leader sequences of a preprotein, or a preproprotein, which is a precursor to a proprotein, having a leader sequence and one or more prosequences, which generally are removed during processing steps that produce active and mature forms of the polypeptide.

A polynucleotide of the present invention may encode a mature or precursor pre-, pro- or prepropolypeptide as discussed above, among others, fused to additional amino acids, such as those which provide additional functionalities. Thus, for instance, the polypeptide may be fused to a marker sequence, such as a peptide, which facilitates purification of the fused polypeptide. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, such as the tag provided in the vector pQE, among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci., USA 86: 821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Typically, it does not adversely affect protein structure or function, and it binds efficiently, selectively and tightly to metal chelate resins, particularly nickel chelate resins. For instance, as is well known, hexa-histidine tags often bind especially well to nickel-NTA resin, which is well known and readily available and can be obtained commercially from, for instance, Qiagen. Moreover, the histidine-metal interaction not only is stable to a variety of conditions useful to remove non-specifically bound material, but also the fusion polypeptide can be bound and removed under mild, non-denaturing conditions. The hexa-histidine tag can be fused most conveniently to the amino or

the carboxyl terminus of the CCII polypeptide. A tag of the hexa-histidine type is particularly useful for bacterial expression.

Another useful marker sequence in certain other preferred embodiments is a hemagglutinin ("HA") tag, particularly when a mammalian cell is used for expression; e.g., COS-7 cells. The HA tag corresponds to an epitope derived of influenza hemagglutinin protein, which has been described by Wilson et al., Cell 37: 767 (1984), for instance.

Deposited materials

A deposit containing a human CCII cDNA has been deposited with the American Type Culture Collection, as noted above. Also as noted above, the human cDNA deposit is referred to herein as "the deposited clone" or as "the cDNA of the deposited clone."

The deposited clone was deposited with the American Type Culture Collection, 12301 Park Lawn Drive, Rockville, Maryland 20852, USA, on January 2, 1996 and assigned ATCC Deposit No. 97405.

The deposited material is a pBluescript SK (-) plasmid (Stratagene, La Jolla, CA) that contains the full length CCII cDNA, referred to as DNA plasmid 951112 upon deposit.

The deposit has been made under the terms of the Budapest Treaty on the international recognition of the deposit of micro-organisms for purposes of patent procedure. The strain will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. The deposit is provided merely as convenience to those of skill in the art and is not an admission that a deposit is required for enablement, such as that required under 35 U.S.C. §112.

The sequence of the polynucleotides contained in the deposited material, as well as the amino acid sequence of the polypeptide encoded thereby, are controlling in the event of any conflict with any description of sequences herein.

A license may be required to make, use or sell the deposited materials, and no such license is hereby granted.

Probes

The present invention further relates to polynucleotides that hybridize to the herein above-described chemokine sequences, particularly CCII sequences. Preferred in this regard are polynucleotides that have at least 70% identity to the sequences described herein above. Particularly preferred are sequences that

have at least 90% identity. Especially preferred are sequences that have at least 95% identity. In this regard, the present invention especially relates to polynucleotides which hybridize under stringent conditions to the herein above-described polynucleotides. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences.

Particularly preferred embodiments in this respect, moreover, are polynucleotides which hybridize to the above-described polynucleotides and encode polypeptides which retain substantially the same biological function or activity as the mature polypeptide encoded by the human cDNA of Figure 1 (SEQ ID NO:1).

As discussed additionally herein regarding polynucleotide assays of the invention, for instance, a probe as discussed above, derived from the full length CCII cDNA, including the entire CCII cDNA of Figure 1 (SEQ ID NO:1), or the coding region of thereof, or any part thereof useful as a probe, may be used as a hybridization probe for cDNA and genomic DNA to isolate full-length cDNAs and genomic clones encoding CCII and to isolate cDNA and genomic clones of other genes that have a high sequence similarity to the human CCII gene. Such probes generally will comprise at least 15 bases. Preferably, such probes will have at least 30 bases and may have at least 50 bases. Particularly preferred probes will have at least 30 bases and will have 50 bases or less.

Such probes may also be used to identify additional cDNA clones corresponding to a full length transcript and a genomic clone or clones that contain the complete human CCII gene including regulatory and promoter regions, exons, and introns.

For example, the coding region of the CCII gene may be isolated by screening using the known DNA sequence to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the present invention is used to screen a library of human cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

The polynucleotides and polypeptides of the present invention may be employed as research reagents and materials for discovery of treatments and diagnostics to human disease, as further discussed herein relating to polynucleotide assays, *inter alia*.

Polypeptides

The present invention further relates to a human CCII polypeptide which has the deduced amino acid sequence of Figure 1 (SEQ ID NO:2).

The invention also relates to fragments, analogs and derivatives of these polypeptides. The terms "fragment," "derivative" and "analog" when referring to the polypeptide of Figure 1 (SEQ ID NO:2), means a polypeptide which retains essentially the same biological function or activity as such polypeptide. Thus, an analog includes a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature polypeptide.

The polypeptide of the present invention may be a recombinant polypeptide, a natural polypeptide or a synthetic polypeptide. In certain preferred embodiments it is a recombinant polypeptide.

The fragment, derivative or analog of the polypeptide of Figure 1 (SEQ ID NO:2) may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

Among the particularly preferred embodiments of the invention in this regard are polypeptides having the amino acid sequence of CCII set out in Figure 1 (SEQ ID NO:2), variants, analogs, derivatives and fragments thereof, and variants, analogs and derivatives of the fragments. Alternatively, particularly preferred embodiments of the invention in this regard are polypeptides having the amino acid sequence of the CCII of the cDNA in the deposited clone, variants, analogs, derivatives and

fragments thereof, and variants, analogs and derivatives of the fragments.

Further particularly preferred in this regard are variants, analogs, derivatives and fragments, and variants, analogs and derivatives of the fragments, having the amino acid sequence of the CCII polypeptide of Figure 1 (SEQ ID NO:2), in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, deleted or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the CCII. Also especially preferred in this regard are conservative substitutions. Most highly preferred are polypeptides having the amino acid sequence of Figure 1 (SEQ ID NO:2) without substitutions.

The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

The polypeptides of the present invention include the polypeptide of SEQ ID NO:2 (in particular the mature polypeptide) as well as polypeptides which have at least 70% similarity (preferably at least 70% identity) to the polypeptide of SEQ ID NO:2 and more preferably at least 90% similarity (more preferably at least 90% identity) to the polypeptide of SEQ ID NO:2 and still more preferably at least 95% similarity (still more preferably at least 95% identity) to the polypeptide of SEQ ID NO:2 and also include portions of such polypeptides with such portion of the polypeptide generally containing at least 30 amino acids and more preferably at least 50 amino acids.

As known in the art "similarity" between two polypeptides is determined by comparing the amino acid sequence and its conserved amino acid substitutes of one polypeptide to the sequence of a second polypeptide.

Fragments or portions of the polypeptides of the present invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, the fragments may be employed as intermediates for producing the full-length polypeptides. Fragments or portions of the polynucleotides of the present invention may be used to synthesize full-length polynucleotides of the present invention.

As part of or following isolation, such polynucleotides can be joined to other polynucleotides, such as DNAs, for mutagenesis, to form fusion proteins, and for propagation or expression in a host, for instance. The isolated polynucleotides, alone or joined to other polynucleotides such as vectors, can be introduced into host cells, in culture or in whole organisms. Introduced into host cells in culture or in whole organisms, such DNAs still would be isolated, as the term is used herein, because they would not be in their naturally occurring form or environment. Similarly, the polynucleotides and polypeptides may occur in a composition, such as a media formulation, a solution for introduction into cells, a composition or solution for chemical or enzymatic reaction, and the like, which are not naturally compositions, and therein remain isolated polynucleotides or polypeptides within the meaning of that term as it is employed herein.

Fragments

Also among preferred embodiments of this aspect of the present invention are polypeptides comprising fragments of CCII, most particularly fragments of the CCII having the amino acid set out in Figure 1 (SEQ ID NO:2), and fragments of variants and derivatives of the CCII of Figure 1 (SEQ ID NO:2).

In this regard a fragment is a polypeptide having an amino acid sequence that entirely is the same as part but not all of the amino acid sequence of the aforementioned CCII polypeptides and variants or derivatives thereof.

Such fragments may be "free-standing," i.e., not part of or fused to other amino acids or polypeptides, or they may be comprised within a larger polypeptide of which they form a part or region. When comprised within a larger polypeptide, the presently discussed fragments most preferably form a single continuous region. However, several fragments may be comprised within a single larger polypeptide. For instance, certain preferred embodiments relate to a fragment of a CCII polypeptide of the present comprised within a precursor polypeptide designed for expression in a host and having heterologous pre and pro-polypeptide regions fused to the amino terminus of the CCII fragment and an additional region fused to the carboxyl terminus of the fragment. Therefore, fragments in one aspect of the meaning

intended herein, refers to the portion or portions of a fusion polypeptide or fusion protein derived from CCII.

Among preferred fragments of CCII are fragments about 5-15, 10-20, 15-40, 25-50, 35-60, 50-75, 65-80, 65-90, 65-98, 50-98, 75-98 and 90-98 amino acids long.

In this context about includes the particularly recited range and ranges larger or smaller by several, a few, 5, 4, 3, 2 or 1 amino acid at either extreme or at both extremes. For instance, about at least 65-98 amino acids in this context means a polypeptide fragment of at least 65, at least 65 plus or minus several, a few, 5, 4, 3, 2 or 1 amino acid to at least 90 or at least 90 plus or minus several a few, 5, 4, 3, 2 or 1 amino acid residues, i.e., ranges a broad as at least 65 minus several amino acids to at least 90 plus several amino acids to as narrow as at least 65 plus several amino acids to at least 90 minus several amino acids.

Highly, preferred in this regard are the recited ranges plus or minus as many as 5 amino acids at either or at both extremes. Particularly highly preferred are the recited ranges means plus or minus as many as 3 amino acids at either or at both extremes. Especially particularly highly preferred are ranges plus or minus 1 amino acid at either or at both extremes. Most highly preferred of all in this regard are fragments at least 5-15, 10-20, 15-40, 25-50, 35-60, 50-75, 65-80, 65-90, 65-98, 50-98, 75-98, and 90-98 amino acids long are preferred.

Among especially preferred fragments of the invention are truncation mutants of CCII. Truncation mutants include CCII polypeptides having the amino acid sequence of Figure 1 (SEQ ID NO:2), or of variants or derivatives thereof, except for deletion of a continuous series of residues (that is, a continuous region, part or portion) that includes the amino terminus, or a continuous series of residues that includes the carboxyl terminus or, as in double truncation mutants, deletion of two continuous series of residues, one including the amino terminus and one including the carboxyl terminus. Fragments having the size ranges set out about also are preferred embodiments of truncation fragments, which are especially preferred among fragments generally.

Also preferred in this aspect of the invention are fragments characterized by structural or functional attributes of CCII.

Preferred embodiments of the invention in this regard include fragments that comprise alpha-helix and alpha-helix forming regions ("alpha-regions"), beta-sheet and beta-sheet-forming regions ("beta-regions"), turn and turn-forming regions ("turn-regions"), coil and coil-forming regions ("coil-regions"), hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions and high antigenic index regions of CCII.

Certain preferred regions in these regards are set out in Figure 3, and include, but are not limited to, regions of the aforementioned types identified by analysis of the amino acid sequence set out in Figure 1 (SEQ ID NO:2). As set out in Figure 3, such preferred regions include Garnier-Robson alpha-regions, beta-regions, turn-regions and coil-regions, Chou-Fasman alpha-regions, beta-regions and turn-regions, Kyte-Doolittle hydrophilic regions and hydrophilic regions, Eisenberg alpha and beta amphipathic regions, Karplus-Schulz flexible regions, Emini surface-forming regions and Jameson-Wolf high antigenic index regions.

Among highly preferred fragments in this regard are those that comprise regions of CCII that combine several structural features, such as several of the features set out above. In this regard, the regions defined by the residues about 20 to about 40, expecially 25 to 35, and about 50 to about 65, especially 56 to 62 and about 65 to about 75, especially 66-70 and about 85 to about 98 of Figure 1 (SEQ ID NO:2), which all are characterized by amino acid compositions highly characteristic of turn-regions, hydrophilic regions, flexible-regions, surface-forming regions, and high antigenic index-regions, are especially highly preferred regions. Such regions may be comprised within a larger polypeptide or may be by themselves a preferred fragment of the present invention, as discussed above. It will be appreciated that the term "about" as used in this paragraph has the meaning set out above regarding fragments in general.

Further preferred regions are those that mediate activities of CCII. Most highly preferred in this regard are fragments that have a chemical, biological or other activity of CCII, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Highly preferred in this regard

are fragments that contain regions that are homologs in sequence, or in position, or in both sequence and to active regions of related polypeptides, such as the related polypeptide set out in Figure 2 (SEQ ID NO:9). Among particularly preferred fragments in these regards are truncation mutants, as discussed above.

It will be appreciated that the invention also relates to, among others, polynucleotides encoding the aforementioned fragments, polynucleotides that hybridize to polynucleotides encoding the fragments, particularly those that hybridize under stringent conditions, and polynucleotides, such as PCR primers, for amplifying polynucleotides that encode the fragments. In these regards, preferred polynucleotides are those that correspond to the preferred fragments, as discussed above.

Vectors, host cells, expression

The present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques.

Host cells can be genetically engineered to incorporate polynucleotides and express polypeptides of the present invention. For instance, polynucleotides may be introduced into host cells using well known techniques of infection, transduction, transfection, transvection and transformation. The polynucleotides may be introduced alone or with other polynucleotides. Such other polynucleotides may be introduced independently, co-introduced or introduced joined to the polynucleotides of the invention.

Thus, for instance, polynucleotides of the invention may be transfected into host cells with another, separate, polynucleotide encoding a selectable marker, using standard techniques for co-transfection and selection in, for instance, mammalian cells. In this case the polynucleotides generally will be stably incorporated into the host cell genome.

Alternatively, the polynucleotides may be joined to a vector containing selectable marker for propagation in a host. The vector construct may be introduced into host cells by the aforementioned techniques. Generally, a plasmid vector is introduced as DNA in a precipitate, such as a calcium phosphate

precipitate, or in a complex with a charged lipid. Electroporation also may be used to introduce polynucleotides into a host. If the vector is a virus, it may be packaged in vitro or introduced into a packaging cell and the packaged virus may be transduced into cells. A wide variety of techniques suitable for making polynucleotides and for introducing polynucleotides into cells in accordance with this aspect of the invention are well known and routine to those of skill in the art. Such techniques are reviewed at length in Sambrook et al. cited above, which is illustrative of the many laboratory manuals that detail these techniques. In accordance with this aspect of the invention the vector may be, for example, a plasmid vector, a single or double-stranded phage vector, a single or double-stranded RNA or DNA viral vector. Such vectors may be introduced into cells as polynucleotides, preferably DNA, by well known techniques for introducing DNA and RNA into cells. The vectors, in the case of phage and viral vectors also may be and preferably are introduced into cells as packaged or encapsidated virus by well known techniques for infection and transduction. Viral vectors may be replication competent or replication defective. In the latter case viral propagation generally will occur only in complementing host cells.

Preferred among vectors, in certain respects, are those for expression of polynucleotides and polypeptides of the present invention. Generally, such vectors comprise cis-acting control regions effective for expression in a host operatively linked to the polynucleotide to be expressed. Appropriate trans-acting factors either are supplied by the host, supplied by a complementing vector or supplied by the vector itself upon introduction into the host.

In certain preferred embodiments in this regard, the vectors provide for specific expression. Such specific expression may be inducible expression or expression only in certain types of cells or both inducible and cell-specific. Particularly preferred among inducible vectors are vectors that can be induced for expression by environmental factors that are easy to manipulate, such as temperature and nutrient additives. A variety of vectors suitable to this aspect of the invention, including constitutive and inducible expression vectors for use in prokaryotic and eukaryotic

hosts, are well known and employed routinely by those of skill in the art.

The engineered host cells can be cultured in conventional nutrient media, which may be modified as appropriate for, inter alia, activating promoters, selecting transformants or amplifying genes. Culture conditions, such as temperature, pH and the like, previously used with the host cell selected for expression generally will be suitable for expression of polypeptides of the present invention as will be apparent to those of skill in the art.

A great variety of expression vectors can be used to express a polypeptide of the invention. Such vectors include chromosomal, episomal and virus-derived vectors e.g., vectors derived from bacterial plasmids, from bacteriophage, from yeast episomes, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids, all may be used for expression in accordance with this aspect of the present invention. Generally, any vector suitable to maintain, propagate or express polynucleotides to express a polypeptide in a host may be used for expression in this regard.

The appropriate DNA sequence may be inserted into the vector by any of a variety of well-known and routine techniques. In general, a DNA sequence for expression is joined to an expression vector by cleaving the DNA sequence and the expression vector with one or more restriction endonucleases and then joining the restriction fragments together using T4 DNA ligase. Procedures for restriction and ligation that can be used to this end are well known and routine to those of skill. Suitable procedures in this regard, and for constructing expression vectors using alternative techniques, which also are well known and routine to those skill, are set forth in great detail in Sambrook et al. cited elsewhere herein.

The DNA sequence in the expression vector is operatively linked to appropriate expression control sequence(s), including, for instance, a promoter to direct mRNA transcription.

Representatives of such promoters include the phage lambda PL promoter, the E. coli lac, trp and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name just a few of the well-known promoters. It will be understood that numerous promoters not mentioned are suitable for use in this aspect of the invention are well known and readily may be employed by those of skill in the manner illustrated by the discussion and the examples herein.

In general, expression constructs will contain sites for transcription initiation and termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will include a translation initiating AUG at the beginning and a termination codon appropriately positioned at the end of the polypeptide to be translated.

In addition, the constructs may contain control regions that regulate as well as engender expression. Generally, in accordance with many commonly practiced procedures, such regions will operate by controlling transcription, such as repressor binding sites and enhancers, among others.

Vectors for propagation and expression generally will include selectable markers. Such markers also may be suitable for amplification or the vectors may contain additional markers for this purpose. In this regard, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells. Preferred markers include dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, and tetracycline or ampicillin resistance genes for culturing E. coli and other bacteria.

The vector containing the appropriate DNA sequence as described elsewhere herein, as well as an appropriate promoter, and other appropriate control sequences, may be introduced into an appropriate host using a variety of well known techniques suitable to expression therein of a desired polypeptide. Representative examples of appropriate hosts include bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS and Bowes melanoma cells; and plant cells. Hosts for of a great variety of

expression constructs are well known, and those of skill will be enabled by the present disclosure readily to select a host for expressing a polypeptides in accordance with this aspect of the present invention.

More particularly, the present invention also includes recombinant constructs, such as expression constructs, comprising one or more of the sequences described above. The constructs comprise a vector, such as a plasmid or viral vector, into which such a sequence of the invention has been inserted. The sequence may be inserted in a forward or reverse orientation. In certain preferred embodiments in this regard, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and there are many commercially available vectors suitable for use in the present invention.

The following vectors, which are commercially available, are provided by way of example. Among vectors preferred for use in bacteria are pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. These vectors are listed solely by way of illustration of the many commercially available and well known vectors that are available to those of skill in the art for use in accordance with this aspect of the present invention. It will be appreciated that any other plasmid or vector suitable for, for example, introduction, maintenance, propagation or expression of a polynucleotide or polypeptide of the invention in a host may be used in this aspect of the invention.

Promoter regions can be selected from any desired gene using vectors that contain a reporter transcription unit lacking a promoter region, such as a chloramphenicol acetyl transferase ("cat") transcription unit, downstream of restriction site or sites for introducing a candidate promoter fragment; i.e., a fragment that may contain a promoter. As is well known, introduction into the vector of a promoter-containing fragment at the restriction

site upstream of the cat gene engenders production of CAT activity, which can be detected by standard CAT assays. Vectors suitable to this end are well known and readily available. Two such vectors are pKK232-8 and pCM7. Thus, promoters for expression of polynucleotides of the present invention include not only well known and readily available promoters, but also promoters that readily may be obtained by the foregoing technique, using a reporter gene.

Among known bacterial promoters suitable for expression of polynucleotides and polypeptides in accordance with the present invention are the E. coli lacI and lacZ and promoters, the T3 and T7 promoters, the gpt promoter, the lambda PR, PL promoters and the trp promoter. Among known eukaryotic promoters suitable in this regard are the CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus ("RSV"), and metallothionein promoters, such as the mouse metallothionein-I promoter.

Selection of appropriate vectors and promoters for expression in a host cell is a well known procedure and the requisite techniques for expression vector construction, introduction of the vector into the host and expression in the host are routine skills in the art.

The present invention also relates to host cells containing the above-described constructs discussed above. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell.

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al. BASIC METHODS IN MOLECULAR BIOLOGY, (1986).

Constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989).

Generally, recombinant expression vectors will include origins of replication, a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence, and a selectable marker to permit isolation of vector containing cells after exposure to the vector. Among suitable promoters are those derived from the genes that encode glycolytic enzymes such as 3-phosphoglycerate kinase ("PGK"), a-factor, acid phosphatase, and heat shock proteins, among others. Selectable markers include the ampicillin resistance gene of E. coli and the trp1 gene of S. cerevisiae.

Transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act to increase transcriptional activity of a promoter in a given host cell-type. Examples of enhancers include the SV40 enhancer, which is located on the late side of the replication origin at bp 100 to 270, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Polynucleotides of the invention, encoding the heterologous structural sequence of a polypeptide of the invention generally will be inserted into the vector using standard techniques so that it is operably linked to the promoter for expression. The polynucleotide will be positioned so that the transcription start site is located appropriately 5' to a ribosome binding site. The ribosome binding site will be 5' to the AUG that initiates translation of the polypeptide to be expressed. Generally, there will be no other open reading frames that begin with an initiation codon, usually AUG, and lie between the ribosome binding site and

the initiating AUG. Also, generally, there will be a translation stop codon at the end of the polypeptide and there will be a polyadenylation signal and a transcription termination signal appropriately disposed at the 3' end of the transcribed region.

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. The signals may be endogenous to the polypeptide or they may be heterologous signals.

The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals but also additional heterologous functional regions. Thus, for instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell, during purification or during subsequent handling and storage. Also, region also may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate purification, among others, are familiar and routine techniques in the art.

Suitable prokaryotic hosts for propagation, maintenance or expression of polynucleotides and polypeptides in accordance with the invention include Escherichia coli, Bacillus subtilis and Salmonella typhimurium. Various species of Pseudomonas, Streptomyces, and Staphylococcus are suitable hosts in this regard. Moreover, many other hosts also known to those of skill may be employed in this regard.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, where the selected promoter is inducible it is induced by appropriate means (e.g., temperature shift or exposure to chemical inducer) and cells are cultured for an additional period.

Cells typically then are harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, such methods are well know to those skilled in the art.

Various mammalian cell culture systems can be employed for expression, as well. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblast, described in Gluzman et al., Cell 23: 175 (1981). Other cell lines capable of expressing a compatible vector include for example, the C127, 3T3, CHO, HeLa, human kidney 293 and BHK cell lines.

Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation sites, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking non-transcribed sequences that are necessary for expression. In certain preferred embodiments in this regard DNA sequences derived from the SV40 splice sites, and the SV40 polyadenylation sites are used for required non-transcribed genetic elements of these types.

The CCII polypeptide can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification. Well known techniques for refolding protein may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and

products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes.

CCII polynucleotides and polypeptides may be used in accordance with the present invention for a variety of applications, particularly those that make use of the chemical and biological properties CCII. Additional applications relate to diagnosis and to treatment of disorders of cells, tissues and organisms. These aspects of the invention are illustrated further by the following discussion.

Polynucleotide assays

This invention is also related to the use of the CCII polynucleotides to detect complementary polynucleotides such as, for example, as a diagnostic reagent. Detection of a mutated form of CCII associated with a dysfunction will provide a diagnostic tool that can add or define a diagnosis of a disease or susceptibility to a disease which results from under-expression over-expression or altered expression of CCII.

Individuals carrying mutations in the human CCII gene may be detected at the DNA level by a variety of techniques. Nucleic acids for diagnosis may be obtained from a patient's cells, such as from blood, urine, saliva, tissue biopsy and autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR prior to analysis. PCR (Saiki et al., Nature, 324: 163-166 (1986)). RNA or cDNA may also be used in the same ways. As an example, PCR primers complementary to the nucleic acid encoding CCII can be used to identify and analyze CCII expression and mutations. For example, deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to radiolabeled CCII RNA or alternatively, radiolabeled CCII antisense DNA sequences. Perfectly matched sequences can be distinguished from mismatched

duplexes by RNase A digestion or by differences in melting temperatures.

Sequence differences between a reference gene and genes having mutations also may be revealed by direct DNA sequencing. In addition, cloned DNA segments may be employed as probes to detect specific DNA segments. The sensitivity of such methods can be greatly enhanced by appropriate use of PCR or another amplification method. For example, a sequencing primer is used with double-stranded PCR product or a single-stranded template molecule generated by a modified PCR. The sequence determination is performed by conventional procedures with radiolabeled nucleotide or by automatic sequencing procedures with fluorescent-tags.

Genetic testing based on DNA sequence differences may be achieved by detection of alteration in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents. Small sequence deletions and insertions can be visualized by high resolution gel electrophoresis. DNA fragments of different sequences may be distinguished on denaturing formamide gradient gels in which the mobilities of different DNA fragments are retarded in the gel at different positions according to their specific melting or partial melting temperatures (see, e.g., Myers et al., Science, 230: 1242 (1985)).

Sequence changes at specific locations also may be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (e.g., Cotton et al., Proc. Natl. Acad. Sci., USA, 85: 4397-4401 (1985)).

Thus, the detection of a specific DNA sequence may be achieved by methods such as hybridization, RNase protection, chemical cleavage, direct DNA sequencing or the use of restriction enzymes, (e.g., restriction fragment length polymorphisms ("RFLP") and Southern blotting of genomic DNA.

In addition to more conventional gel-electrophoresis and DNA sequencing, mutations also can be detected by in situ analysis.

Chromosome assays

The sequences of the present invention are also valuable for chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Moreover, there is a current need for

identifying particular sites on the chromosome. Few chromosome marking reagents based on actual sequence data (repeat polymorphisms) are presently available for marking chromosomal location. The mapping of DNAs to chromosomes according to the present invention is an important first step in correlating those sequences with genes associated with disease.

In certain preferred embodiments in this regard, the cDNA herein disclosed is used to clone genomic DNA of a CCII gene. This can be accomplished using a variety of well known techniques and libraries, which generally are available commercially. The genomic DNA the is used for in situ chromosome mapping using well known techniques for this purpose. Typically, in accordance with routine procedures for chromosome mapping, some trial and error may be necessary to identify a genomic probe that gives a good in situ hybridization signal.

In some cases, in addition, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the cDNA. Computer analysis of the 3' untranslated region of the gene is used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the primer will yield an amplified fragment.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular DNA to a particular chromosome. Using the present invention with the same oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes or pools of large genomic clones in an analogous manner. Other mapping strategies that can similarly be used to map to its chromosome include in situ hybridization, prescreening with labeled flow-sorted chromosomes and preselection by hybridization to construct chromosome specific-cDNA libraries.

Fluorescence in situ hybridization (FISH) of a cDNA clone to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with cDNA as short as 50 or 60. For a review of this technique, see Verma et al., HUMAN CHROMOSOMES: A MANUAL OF BASIC TECHNIQUES, Pergamon Press, New York (1988).

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in V. McKusick, MENDELIAN INHERITANCE IN MAN, available on line through Johns Hopkins University, Welch Medical Library. The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

Next, it is necessary to determine the differences in the cDNA or genomic sequence between affected and unaffected individuals. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

With current resolution of physical mapping and genetic mapping techniques, a cDNA precisely localized to a chromosomal region associated with the disease could be one of between 50 and 500 potential causative genes. (This assumes 1 megabase mapping resolution and one gene per 20 kb).

Polypeptide assays

The present invention also relates to a diagnostic assays such as quantitative and diagnostic assays for detecting levels of CCII protein in cells and tissues, including determination of normal and abnormal levels. Thus, for instance, a diagnostic assay in accordance with the invention for detecting over-expression of CCII protein compared to normal control tissue samples may be used to detect the presence of an immune disorder, for example. Assay techniques that can be used to determine levels of a protein, such as an CCII protein of the present invention, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays. Among these ELISAs frequently are preferred. An ELISA assay initially comprises preparing an antibody specific to CCII, preferably a monoclonal antibody. In addition a reporter antibody generally is prepared which binds to the monoclonal antibody. The reporter antibody is attached a detectable reagent such as radioactive, fluorescent or enzymatic reagent, in this example horseradish peroxidase enzyme.

To carry out an ELISA a sample is removed from a host and incubated on a solid support, e.g. a polystyrene dish, that binds the proteins in the sample. Any free protein binding sites on the dish are then covered by incubating with a non-specific protein such as bovine serum albumin. Next, the monoclonal antibody is incubated in the dish during which time the monoclonal antibodies attach to any CCII proteins attached to the polystyrene dish. Unbound monoclonal antibody is washed out with buffer. The reporter antibody linked to horseradish peroxidase is placed in the dish resulting in binding of the reporter antibody to any monoclonal antibody bound to CCII. Unattached reporter antibody is then washed out. Reagents for peroxidase activity, including a colorimetric substrate are then added to the dish. Immobilized peroxidase, linked to CCII through the primary and secondary antibodies, produces a colored reaction product. The amount of color developed in a given time period indicates the amount of CCII protein present in the sample. Quantitative results typically are obtained by reference to a standard curve.

A competition assay may be employed wherein antibodies specific to CCII attached to a solid support and labeled CCII and a sample derived from the host are passed over the solid support and the amount of label detected attached to the solid support can be correlated to a quantity of CCII in the sample.

Immunoassays and reagents

The polypeptides, their fragments or other derivatives, or analogs thereof, or cells expressing them can be used as an immunogen to produce antibodies thereto. These antibodies can be, for example, polyclonal or monoclonal antibodies. The present invention also includes chimeric, single chain, and humanized antibodies, as well as Fab fragments, or the product of an Fab expression library. Various procedures known in the art may be used for the production of such antibodies and fragments.

Antibodies generated against the polypeptides corresponding to a sequence of the present invention can be obtained by direct injection of the polypeptides into an animal or by administering the polypeptides to an animal, preferably a nonhuman. The antibody so obtained will then bind the polypeptides itself. In this manner, even a sequence encoding only a fragment of the

polypeptides can be used to generate antibodies binding the whole native polypeptides. Such antibodies can then be used to isolate the polypeptide from tissue expressing that polypeptide.

For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., Nature 256: 495-497 (1975), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., Immunology Today 4: 72 (1983) and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., pg. 77-96 in MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc. (1985).

Techniques described for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to immunogenic polypeptide products of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies to immunogenic polypeptide products of this invention.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or purify the polypeptide of the present invention by attachment of the antibody to a solid support for isolation and/or purification by affinity chromatography.

Thus, among others, the chemotactic cellular immune activities of CCII is useful to treat tumors, chronic infections, leukemia, T-cell mediated auto-immune diseases, parasitic infections, psoriasis, asthma, allergy, to regulate hematopoiesis, to stimulate growth factor activity, to inhibit angiogenesis, to promote wound healing, to treat inflammatory disorders, to control cellular immune reactions, to treat malignant diseases, and to inhibit casein kinase to activity.

More specifically, CCII polypeptides may be employed to inhibit bone marrow stem cell colony formation as adjunct protective treatment during cancer chemotherapy and for leukemia.

CCII polypeptides may also be employed to inhibit epidermal keratinocyte proliferation for treatment of psoriasis, which is characterized by keratinocyte hyper-proliferation.

CCII polypeptides may also be employed to treat solid tumors by stimulating the invasion and activation of host defense cells,

e.g., cytotoxic T cells and macrophages and by inhibiting the angiogenesis of tumors. They may also be employed to enhance host defenses against resistant chronic and acute infections, for example, mycobacterial infections via the attraction and activation of microbicidal leukocytes.

CCII polypeptides may also be employed to inhibit T cell proliferation by the inhibition of IL-2 biosynthesis for the treatment of T-cell mediated auto-immune diseases and lymphocytic leukemias.

CCII polypeptides may also be employed to stimulate wound healing, both via the recruitment of debris clearing and connective tissue promoting inflammatory cells and also via its control of excessive TGFβ-mediated fibrosis. In this same manner, 1 may also be employed to treat other fibrotic disorders, including liver cirrhosis, osteoarthritis and pulmonary fibrosis.

CCII polypeptides of the present invention may also be employed as cytostatic agents for antibacterial and antimicrobial functions.

They may also be employed to regulate hematopoiesis, by regulating the activation and differentiation of various hematopoietic progenitor cells, for example, to release mature leukocytes from the bone marrow following chemotherapy.

In these regards, CCII polypeptides are preferred, particularly the CCII having the amino acid sequence set out in Figure 1 (SEQ ID NO:2).

As set out further below, these and other activities and properties of the CCII polynucleotides and polypeptides of the invention have various applications and uses in numerous fields including applications involving chemotaxis and cellular immune reactions.

CCII binding molecules and assays

This invention also provides a method for identification of molecules, such as receptor molecules, that bind CCII. Genes encoding proteins that bind CCII, such as receptor proteins, can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Such methods are described in many laboratory manuals such as, for instance,

Coligan et al., Current Protocols in Immunology 1(2): Chapter 5 (1991).

For instance, expression cloning may be employed for this purpose. To this end polyadenylated RNA is prepared from a cell responsive to CCII, a cDNA library is created from this RNA, the library is divided into pools and the pools are transfected individually into cells that are not responsive to CCII. The transfected cells then are exposed to labeled CCII. (CCII can be labeled by a variety of well-known techniques including standard methods of radio-iodination or inclusion of a recognition site for a site-specific protein kinase.) Following exposure, the cells are fixed and binding of CCII is determined. These procedures conveniently are carried out on glass slides.

Pools are identified of cDNA that produced CCII-binding cells. Sub-pools are prepared from these positives, transfected into host cells and screened as described above. Using an iterative sub-pooling and re-screening process, one or more single clones that encode the putative binding molecule, such as a receptor molecule, can be isolated.

Alternatively a labeled ligand can be photoaffinity linked to a cell extract, such as a membrane or a membrane extract, prepared from cells that express a molecule that it binds, such as a receptor molecule. Cross-linked material is resolved by polyacrylamide gel electrophoresis ("PAGE") and exposed to X-ray film. The labeled complex containing the ligand-receptor can be excised, resolved into peptide fragments, and subjected to protein microsequencing. The amino acid sequence obtained from microsequencing can be used to design unique or degenerate oligonucleotide probes to screen cDNA libraries to identify genes encoding the putative receptor molecule.

Polypeptides of the invention also can be used to assess CCII binding capacity of CCII binding molecules, such as receptor molecules, in cells or in cell-free preparations.

Agonists and antagonists - assays and molecules

The invention also provides a method of screening compounds to identify those which enhance or block the action of CCII on cells, such as its interaction with CCII-binding molecules such as receptor molecules. An agonist is a compound which increases the

natural biological functions of CCII or which functions in a manner similar to CCII, while antagonists decrease or eliminate such functions.

For example, a cellular compartment, such as a membrane or a preparation thereof, such as a membrane-preparation, may be prepared from a cell that expresses a molecule that binds CCII, such as a molecule of a signaling or regulatory pathway modulated by CCII. The preparation is incubated with labeled CCII in the absence or the presence of a candidate molecule which may be a CCII agonist or antagonist. The ability of the candidate molecule to bind the binding molecule is reflected in decreased binding of the labeled ligand. Molecules which bind gratuitously, i.e., without inducing the effects of CCII on binding the CCII binding molecule, are most likely to be good antagonists. Molecules that bind well and elicit effects that are the same as or closely related to CCII are agonists.

CCII-like effects of potential agonists and antagonists may by measured, for instance, by determining activity of a second messenger system following interaction of the candidate molecule with a cell or appropriate cell preparation, and comparing the effect with that of CCII or molecules that elicit the same effects as CCII. Second messenger systems that may be useful in this regard include but are not limited to AMP guanylate cyclase, ion channel or phosphoinositide hydrolysis second messenger systems.

Another example of an assay for CCII antagonists is a competitive assay that combines CCII and a potential antagonist with membrane-bound CCII receptor molecules or recombinant CCII receptor molecules under appropriate conditions for a competitive inhibition assay. CCII can be labeled, such as by radioactivity, such that the number of CCII molecules bound to a receptor molecule can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a polypeptide of the invention and thereby inhibit or extinguish its activity. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds the same sites on a binding molecule, such as

a receptor molecule, without inducing CCII-induced activities, thereby preventing the action of CCII by excluding CCII from binding.

Potential antagonists include a small molecule which binds to and occupies the binding site of the polypeptide thereby preventing binding to cellular binding molecules, such as receptor molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules.

Other potential antagonists include antisense molecules. Antisense technology can be used to control gene expression through antisense DNA or RNA or through triple-helix formation. Antisense techniques are discussed, for example, in - Okano, J. Neurochem. 56: 560 (1991); OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION, CRC Press, Boca Raton, FL (1988). Triple helix formation is discussed in, for instance Lee et al., Nucleic Acids Research 6: 3073 (1979); Cooney et al., Science 241: 456 (1988); and Dervan et al., Science 251: 1360 (1991). The methods are based on binding of a polynucleotide to a complementary DNA or RNA. For example, the 5' coding portion of a polynucleotide that encodes the mature polypeptide of the present invention may be used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription thereby preventing transcription and the production of CCII. The antisense RNA oligonucleotide hybridizes to the mRNA in vivo and blocks translation of the mRNA molecule into CCII polypeptide. The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed in vivo to inhibit production of CCII.

The antagonists may be employed in a composition with a pharmaceutically acceptable carrier, e.g., as hereinafter described.

The antagonists may be employed for instance to treat certain auto-immune diseases, atherosclerosis, chronic inflammatory and infectious diseases, histamine and IgE-mediated allergic reactions, prostaglandin-independent fever, bone marrow failure, cancers, silicosis, sarcoidosis, rheumatoid arthritis, shock, hyper-eosinophilic syndrome and fibrosis in the asthmatic lung, cystic

fibrosis, malignant diseases, psoriasis, diapedesis and urinary and kidney stones.

Compositions

The invention also relates to compositions comprising the polynucleotide or the polypeptides discussed above or the agonists or antagonists. Thus, the polypeptides of the present invention may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to a subject. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a polypeptide of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration.

Kits

The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, reflecting approval by the agency of the manufacture, use or sale of the product for human administration.

Administration

Polypeptides of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

The pharmaceutical compositions generally are administered in an amount effective for treatment or prophylaxis of a specific indication or indications. In general, the compositions are

administered in an amount of at least about 10 μg/kg body weight. In most cases they will be administered in an amount not in excess of about 8 mg/kg body weight per day. Preferably, in most cases, dose is from about 10 μg/kg to about 1 mg/kg body weight, daily. It will be appreciated that optimum dosage will be determined by standard methods for each treatment modality and indication, taking into account the indication, its severity, route of administration, complicating conditions and the like.

Gene therapy

The CCII polynucleotides, polypeptides, agonists and antagonists that are polypeptides may be employed in accordance with the present invention by expression of such polypeptides in vivo, in treatment modalities often referred to as "gene therapy."

Thus, for example, cells from a patient may be engineered with a polynucleotide, such as a DNA or RNA, encoding a polypeptide ex vivo, and the engineered cells then can be provided to a patient to be treated with the polypeptide. For example, cells may be engineered ex vivo by the use of a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention. Such methods are well-known in the art and their use in the present invention will be apparent from the teachings herein.

Similarly, cells may be engineered in vivo for expression of a polypeptide in vivo by procedures known in the art. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector, as discussed above. The retroviral expression construct then may be isolated and introduced into a packaging cell is transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a patient for engineering cells in vivo and expression of the polypeptide in vivo. These and other methods for administering a polypeptide of the present invention by such method should be apparent to those skilled in the art from the teachings of the present invention.

Retroviruses from which the retroviral plasmid vectors herein above mentioned may be derived include, but are not limited to, Moloney Murine Leukemia Virus, spleen necrosis virus, retroviruses

such as Rous Sarcoma Virus, Harvey Sarcoma Virus, avian leukosis virus, gibbon ape leukemia virus, human immunodeficiency virus, adenovirus, Myeloproliferative Sarcoma Virus, and mammary tumor virus. In one embodiment, the retroviral plasmid vector is derived from Moloney Murine Leukemia Virus.

Such vectors well include one or more promoters for expressing the polypeptide. Suitable promoters which may be employed include, but are not limited to, the retroviral LTR; the SV40 promoter; and the human cytomegalovirus (CMV) promoter described in Miller et al., Biotechniques 7: 980-990 (1989), or any other promoter (e.g., cellular promoters such as eukaryotic cellular promoters including, but not limited to, the histone, RNA polymerase III, and ß-actin promoters). Other viral promoters which may be employed include, but are not limited to, adenovirus promoters, thymidine kinase (TK) promoters, and B19 parvovirus promoters. The selection of a suitable promoter will be apparent to those skilled in the art from the teachings contained herein.

The nucleic acid sequence encoding the polypeptide of the present invention will be placed under the control of a suitable promoter. Suitable promoters which may be employed include, but are not limited to, adenoviral promoters, such as the adenoviral major late promoter; or heterologous promoters, such as the cytomegalovirus (CMV) promoter; the respiratory syncytial virus (RSV) promoter; inducible promoters, such as the MMT promoter, the metallothionein promoter; heat shock promoters; the albumin promoter; the ApoAI promoter; human globin promoters; viral thymidine kinase promoters, such as the Herpes Simplex thymidine kinase promoter; retroviral LTRs (including the modified retroviral LTRs herein above described); the ß-actin promoter; and human growth hormone promoters. The promoter also may be the native promoter which controls the gene encoding the polypeptide.

The retroviral plasmid vector is employed to transduce packaging cell lines to form producer cell lines. Examples of packaging cells which may be transfected include, but are not limited to, the PE501, PA317, Y-2, Y-AM, PA12, T19-14X, VT-19-17-H2, YCRE, YCRIP, GP+E-86, GP+envAm12, and DAN cell lines as described in Miller, A., Human Gene Therapy 1: 5-14 (1990). The vector may be transduced into the packaging cells through any means known in the art. Such means include, but are not limited to,

electroporation, the use of liposomes, and CaPO4 precipitation. In one alternative, the retroviral plasmid vector may be encapsulated into a liposome, or coupled to a lipid, and then administered to a host.

The producer cell line will generate infectious retroviral vector particles, which include the nucleic acid sequence(s) encoding the polypeptides. Such retroviral vector particles then may be employed to transduce eukaryotic cells, either in vitro or in vivo. The transduced eukaryotic cells will express the nucleic acid sequence(s) encoding the polypeptide. Eukaryotic cells which may be transduced include, but are not limited to, embryonic stem cells, embryonic carcinoma cells, as well as hematopoietic stem cells, hepatocytes, fibroblasts, myoblasts, keratinocytes, endothelial cells, and bronchial epithelial cells.

EXAMPLES

The present invention is further described by the following examples. The examples are provided solely to illustrate the invention by reference to specific embodiments. These exemplifications, while illustrating certain specific aspects of the invention, do not portray the limitations or circumscribe the scope of the disclosed invention.

Certain terms used herein are explained in the foregoing glossary.

All examples were carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. Routine molecular biology techniques of the following examples can be carried out as described in standard laboratory manuals, such as Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989), herein referred to as "Sambrook."

All parts or amounts set out in the following examples are by weight, unless otherwise specified.

Unless otherwise stated size separation of fragments in the examples below was carried out using standard techniques of agarose and polyacrylamide gel electrophoresis ("PAGE") in Sambrook and numerous other references such as, for instance, by Goeddel et al., Nucleic Acids Res. 8: 4057 (1980).

Unless described otherwise, ligations were accomplished using standard buffers, incubation temperatures and times, approximately equimolar amounts of the DNA fragments to be ligated and approximately 10 units of T4 DNA ligase ("ligase") per 0.5 μg of DNA.

Example 1 Expression and purification of human CCII using bacteria

The DNA sequence encoding human CCII in the deposited polynucleotide was amplified using PCR oligonucleotide primers specific to the amino acid carboxyl terminal sequence of the human CCII protein and to vector sequences 3' to the gene. Additional nucleotides containing restriction sites to facilitate cloning were added to the 5' and 3' sequences respectively.

The 5' oligonucleotide primer had the sequence 5' CGCC CATGGCAGCAGAACCACTGA 3' (SEQ ID NO:3) containing the underlined NcoI restriction site, which encodes a start AUG, followed by 16 nucleotides of the human CCII coding sequence set out in Figure 1 (SEQ ID NO:1) beginning with the first base of the second codon.

The 3' primer had the sequence 5' CGC AAG CTT AGCCAGGCGGCTTTA 3' (SEQ ID NO:10) the underlined Hind III restriction site followed by 15 nucleotides complementary to 15 nucleotides of the CCII non-coding sequence set out in Figure 1 (SEQ ID NO:1), including the stop codon.

The restrictions sites were convenient to restriction enzyme sites in the bacterial expression vectors pQE-9, which were used for bacterial expression in these examples. (Qiagen, Inc. Chatsworth, CA). pQE-9 encodes ampicillin antibiotic resistance ("Ampr") and contains a bacterial origin of replication ("ori"), an IPTG inducible promoter, a ribosome binding site ("RBS"), a 6-His tag and restriction enzyme sites.

The amplified human CCII DNA and the vector pQE-9 both were digested with NcoI and HindIII, and the digested DNAs then were ligated together. Insertion of the CCII DNA into the NcoI/HindIII restricted vector placed the CCII coding region downstream of and operably linked to the vector's IPTG-inducible promoter and in-frame with an initiating AUG appropriately positioned for translation of CCII.

The ligation mixture was transformed into competent E. coli cells using standard procedures. Such procedures are described in

Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989). E. coli strain M15/rep4, containing multiple copies of the plasmid pREP4, which expresses lac repressor and confers kanamycin resistance ("Kanr"), was used in carrying out the illustrative example described here. This strain, which is only one of many that are suitable for expressing CCII, is available commercially from Qiagen.

Transformants were identified by their ability to grow on LB plates in the presence of ampicillin. Plasmid DNA was isolated from resistant colonies and the identity of the cloned DNA was confirmed by restriction analysis.

Clones containing the desired constructs were grown overnight ("O/N") in liquid culture in LB media supplemented with both ampicillin (100 ug/ml) and kanamycin (25 ug/ml).

The O/N culture was used to inoculate a large culture, at a dilution of approximately 1:100 to 1:250. The cells were grown to an optical density at 600nm ("OD600") of between 0.4 and 0.6. Isopropyl-B-D-thiogalactopyranoside ("IPTG") was then added to a final concentration of 1 mM to induce transcription from lac repressor sensitive promoters, by inactivating the lacI repressor. Cells subsequently were incubated further for 3 to 4 hours. Cells then were harvested by centrifugation and disrupted, by standard methods. Inclusion bodies were purified from the disrupted cells using routine collection techniques, and protein was solubilized from the inclusion bodies into 8M urea. The 8M urea solution containing the solubilized protein was passed over a PD-10 column in 2X phosphate buffered saline ("PBS"), thereby removing the urea, exchanging the buffer and refolding the protein. The protein was purified by a further step of chromatography to remove endotoxin. Then, it was sterile filtered. The sterile filtered protein preparation was stored in 2X PBS at a concentration of 95 micrograms per mL.

Analysis of the preparation by standard methods of polyacrylamide gel electrophoresis revealed that the preparation contained about 90% monomer CCII having the expected molecular weight of, approximately, 10 kDa.

Example 2 Cloning and expression of human CCII in a baculovirus expression system

The cDNA sequence encoding the full length human CCII protein, in the deposited clone is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene:

The 5' primer has the sequence 5' CGC GGA TCC CGC AGC AGA ACC ACT G 3' (SEQ ID NO:5) containing the underlined BamHI restriction enzyme site followed by 16 bases of the sequence of CCII of Figure 1 (SEQ ID NO:1). Inserted into an expression vector, as described below, the 5' end of the amplified fragment encoding human CCII provides an efficient signal peptide. An efficient signal for initiation of translation in eukaryotic cells, as described by Kozak, M., J. Mol. Biol. 196: 947-950 (1987) is appropriately located in the vector portion of the construct. The 3' primer has the sequence 5' CGC GGT ACC AGC CAG GCG GCT TTA 3' (SEQ ID NO:6) containing the underlined Asp718 restriction site followed by nucleotides complementary to 15 nucleotides of the CCII non-coding sequence set out in Figure 1 (SEQ ID NO:1).

The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Ca.). The fragment then is digested with BamH1 and Asp718 and again is purified on a 1% agarose gel. This fragment is designated herein F2.

The vector pA2-GP is used to express the CCII protein in the baculovirus expression system, using standard methods, such as those described in Summers et al, A MANUAL OF METHODS FOR BACULOVIRUS VECTORS AND INSECT CELL CULTURE PROCEDURES, Texas Agricultural Experimental Station Bulletin No. 1555 (1987). This expression vector contains the strong polyhedrin promoter of the Autographa californica nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites. The signal peptide of AcMNPV gp67, including the N-terminal methionine, is located just upstream of a BamH1 site. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For an easy selection of recombinant virus the beta-galactosidase gene from E.coli is inserted in the same orientation as the polyhedrin promoter and is followed by the polyadenylation signal of the polyhedrin gene. The polyhedrin sequences are flanked at both sides by viral sequences for cell-mediated homologous recombination

with wild-type viral DNA to generate viable virus that express the cloned polynucleotide.

Many other baculovirus vectors could be used in place of pA2-GP, such as pAc373, pVL941 and pAcIM1 provided, as those of skill readily will appreciate, that construction provides appropriately located signals for transcription, translation, trafficking and the like, such as an in-frame AUG and a signal peptide, as required. Such vectors are described in Luckow et al., Virology 170: 31-39, among others.

The plasmid is digested with the restriction enzymes BamH1 and Asp718 and then is dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Ca.). This vector DNA is designated herein "V2".

Fragment F2 and the dephosphorylated plasmid V2 are ligated together with T4 DNA ligase. E.coli HB101 cells are transformed with ligation mix and spread on culture plates. Bacteria are identified that contain the plasmid with the human CCII gene by digesting DNA from individual colonies using BamH1 and Asp718 and then analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing. This plasmid is designated herein pBacCCII.

5 μg of the plasmid pBacCCII is co-transfected with 1.0 μg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA.), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84: 7413-7417 (1987). 1μg of BaculoGold™ virus DNA and 5 μg of the plasmid pBacCCII are mixed in a sterile well of a microtiter plate containing 50 μl of serum free Grace's medium (Life Technologies Inc., Gaithersburg, MD). Afterwards 10 μl Lipofectin plus 90 μl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is rocked back and forth to mix the newly added solution. The plate is then incubated for 5 hours at 27°C. After 5 hours the transfection solution is removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum

is added. The plate is put back into an incubator and cultivation is continued at 27°C for four days.

After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, cited above. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10).

Four days after serial dilution, the virus is added to the cells. After appropriate incubation, blue stained plaques are picked with the tip of an Eppendorf pipette. The agar containing the recombinant viruses is then resuspended in an Eppendorf tube containing 200 µl of Grace's medium. The agar is removed by a brief centrifugation and the supernatant containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4°C. A clone containing properly inserted CCII is identified by DNA analysis including restriction mapping and sequencing. This is designated herein as V-CCII.

Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus V-CCII at a multiplicity of infection ("MOI") of about 2 (about 1 to about 3). Six hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Gaithersburg). 42 hours later, 5 µCi of 35S-methionine and 5 µCi 35S cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then they are harvested by centrifugation, lysed and the labeled proteins are visualized by SDS-PAGE and autoradiography.

Baculovirus infected SF-9 cell supernatant was acidified with acetic acid to pH 5 and passed through a Perseptive Biosystems HS50 cation exchange resin in 20 mM acetic acid/sodium acetate buffer pH 5 with a 20 mM to 1.5 M NaCl salt gradient in the same buffer. Fractions containing protein were confirmed by microsequencing after transfer onto a Problot membrane. Fractions are pooled and

diluted 1:10 into 20 mM acetic acid/acetate buffer pH 5.0 and passed through Perseptive Biosystems CM20 cation exchange resin with a 20 mM to 2 M salt gradient. Fractions containing the protein as judged by SDS-PAGE were pooled, and sized onto a Pharmacia Sepharose 12 column in the same buffer containing 250 mM NaCl.

Example 3 Expression of CCII in COS cells

The expression plasmid, CCII HA, is made by cloning a cDNA encoding CCII into the expression vector pcDNAI/Amp (which can be obtained from Invitrogen, Inc.).

The expression vector pcDNAI/amp contains: (1) an E.coli origin of replication effective for propagation in E. coli and other prokaryotic cell; (2) an ampicillin resistance gene for selection of plasmid-containing prokaryotic cells; (3) an SV40 origin of replication for propagation in eukaryotic cells; (4) a CMV promoter, a polylinker, an SV40 intron, and a polyadenylation signal arranged so that a cDNA conveniently can be placed under expression control of the CMV promoter and operably linked to the SV40 intron and the polyadenylation signal by means of restriction sites in the polylinker.

A DNA fragment encoding the entire CCII precursor and a HA tag fused in frame to its 3' end is cloned into the polylinker region of the vector so that recombinant protein expression is directed by the CMV promoter. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein described by Wilson et al., Cell 37: 767 (1984). The fusion of the HA tag to the target protein allows easy detection of the recombinant protein with an antibody that recognizes the HA epitope.

The plasmid construction strategy is as follows.

The CCII cDNA of the deposit clone is amplified using primers that contained convenient restriction sites, much as described above regarding the construction of expression vectors for expression of CCII in E. coli and S. fugiperda.

To facilitate detection, purification and characterization of the expressed CCII, one of the primers contains a hemagglutinin tag ("HA tag") as described above.

Suitable primers include that following, which are used in this example.

The 5' primer, 5' CGCGGATCCACCATGGCAGCAGAACCA 3' (SEQ ID NO:7) contains the underlined BamHI site, an AUG start codon and 12 codons thereafter. The 3' primer, containing the underlined Xba I site, the hexapeptide hemaglutinin tag (bold) and last 15 bp of 3' coding sequence (at the 3' end) has the following sequence 5' CGCTCTAGATCAAGCGTAGTCTGGGACGTCGTATGGGTACTTCTTCCTGATCTT 3' (SEQ ID NO:8). The PCR amplified DNA fragment and the vector, pcDNAI/Amp, are digested with and then ligated. The ligation mixture is transformed into E. coli strain SURE (Stratagene Cloning Systems, La Jolla, CA) the transformed culture is plated on ampicillin media plates which then are incubated to allow growth of ampicillin resistant colonies. Plasmid DNA is isolated from resistant colonies and examined by restriction analysis and gel sizing for the presence of the CCII-encoding fragment.

For expression of recombinant CCII, COS cells are transfected with an expression vector, as described above, using DEAE-DEXTRAN, as described, for instance, in Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, Cold Spring Laboratory Press, Cold Spring Harbor, New York (1989).

Cells are incubated under conditions for expression of CCII by the vector.

Expression of the CCII HA fusion protein is detected by radiolabelling and immunoprecipitation, using methods described in, for example Harlow et al., ANTIBODIES: A LABORATORY MANUAL, 2nd Ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1988). To this end, two days after transfection, the cells are labeled by incubation in media containing 35S-cysteine for 8 hours. The cells and the media are collected, and the cells are washed and the lysed with detergent-containing RIPA buffer: 150 mM NaCl, 1% NP-40, 0.1% SDS, 1% NP-40, 0.5% DOC, 50 mM TRIS, pH 7.5, as described by Wilson et al. cited above. Proteins are precipitated from the cell lysate and from the culture media using an HA-specific monoclonal antibody. The precipitated proteins then are analyzed by SDS-PAGE gels and autoradiography. An expression product of the expected size is seen in the cell lysate, which is not seen in negative controls.

Example 4 Tissue distribution of CCII expression

Northern blot analysis is carried out to examine the levels of expression of CCII in human tissues, using methods described by, among others, Sambrook et al, cited above. Total cellular RNA samples are isolated with RNAzol™ B system (Biotecx Laboratories, Inc. 6023 South Loop East, Houston, TX 77033).

About 10μg of Total RNA is isolated from tissue samples. The RNA is size resolved by electrophoresis through a 1% agarose gel under strongly denaturing conditions. RNA is blotted from the gel onto a nylon filter, and the filter then is prepared for hybridization to a detectably labeled polynucleotide probe.

As a probe to detect mRNA that encodes CCII, the antisense strand of the coding region of the cDNA insert in the deposited clone is labeled to a high specific activity. The cDNA is labeled by primer extension, using the Prime-It kit, available from Stratagene. The reaction is carried out using 50 ng of the cDNA, following the standard reaction protocol as recommended by the supplier. The labeled polynucleotide is purified away from other labeled reaction components by column chromatography using a Select-G-50 column, obtained from 5-Prime - 3-Prime, Inc. of 5603 Arapahoe Road, Boulder, CO 80303.

The labeled probe is hybridized to the filter, at a concentration of 1,000,000 cpm/ml, in a small volume of 7% SDS, 0.5 M NaPO4, pH 7.4 at 65°C, overnight.

Thereafter the probe solution is drained and the filter is washed twice at room temperature and twice at 60°C with 0.5 x SSC, 0.1% SDS. The filter then is dried and exposed to film at -70°C overnight with an intensifying screen.

Autoradiography shows that mRNA for CCII is abundant in [breast lymph node cells].

Example 5 Gene therapeutic expression of human CCII

Fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature overnight. After 24 hours at room temperature, the flask is inverted - the chunks of tissue remain fixed to the bottom of the flask - and fresh media is added

(e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin). The tissue is then incubated at 37°C for approximately one week. At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerges. The monolayer is trypsinized and scaled into larger flasks.

A vector for gene therapy is digested with restriction enzymes for cloning a fragment to be expressed. The digested vector is treated with calf intestinal phosphatase to prevent self-ligation. The dephosphorylated, linear vector is fractionated on an agarose gel and purified.

CCII cDNA capable of expressing active CCII, is isolated. The ends of the fragment are modified, if necessary, for cloning into the vector. For instance, 5" overhanging may be treated with DNA polymerase to create blunt ends. 3' overhanging ends may be removed using S1 nuclease. Linkers may be ligated to blunt ends with T4 DNA ligase.

Equal quantities of the Moloney murine leukemia virus linear backbone and the CCII fragment are mixed together and joined using T4 DNA ligase. The ligation mixture is used to transform E. Coli and the bacteria are then plated onto agar-containing kanamycin. Kanamycin phenotype and restriction analysis confirm that the vector has the properly inserted gene.

Packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The vector containing the CCII gene is introduced into the packaging cells by standard techniques. Infectious viral particles containing the CCII gene are collected from the packaging cells, which now are called producer cells.

Fresh media is added to the producer cells, and after an appropriate incubation period media is harvested from the plates of confluent producer cells. The media, containing the infectious viral particles, is filtered through a Millipore filter to remove detached producer cells. The filtered media then is used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the filtered media. Polybrene (Aldrich) may be included in the media to facilitate transduction. After appropriate incubation, the media is removed

and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his, to select out transduced cells for expansion.

Engineered fibroblasts then may be injected into rats, either alone or after having been grown to confluence on microcarrier beads, such as cytodex 3 beads. The injected fibroblasts produce CCII product, and the biological actions of the protein are conveyed to the host.

It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples.

Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Ni, Jian

(ii) TITLE OF INVENTION: CHEMOTACTIC CYTOKINE II

(iii) NUMBER OF SEQUENCES: 10

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI, STEWART & OLSTEIN
(B) STREET: 6 BECKER FARM ROAD
(C) CITY: ROSELAND
(D) STATE: NJ
(E) COUNTRY: USA
(F) ZIP: 07068

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Ferraro, Gregory D
(B) REGISTRATION NUMBER: 36,134
(C) REFERENCE/DOCKET NUMBER: 325800-524

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 201-994-1700
(B) TELEFAX: 201-994-1744

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 532 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 145..438

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
CGAGGCAGCT CTCTCCTCCT TCCCCGCTGC TTATAAACCT CAGCCCTGAG GCTCCAGCTC       60

ACTCTACCCC ATCTCCTTGC CGGGTCAGCC CTGACAAAGG TCAGCTAGCC CCTTGAGGAC      120

ATCAGCTTTG GCCTCAGGGT CCTA ATG GCA GCA GAA CCA CTG ACA GAG CTA          171
                           Met Ala Ala Glu Pro Leu Thr Glu Leu
                             1               5

GAG GAG TCC ATT GAG ACC GTG GTC ACC ACC TTC TTC ACC TTT GCA AGG         219
Glu Glu Ser Ile Glu Thr Val Val Thr Thr Phe Phe Thr Phe Ala Arg
 10                  15                  20                  25
```

```
CAG GAG GGC CGG AAG GAT AGC CTC AGC GTC AAC GAG TTC AAA GAG CTG          267
Gln Glu Gly Arg Lys Asp Ser Leu Ser Val Asn Glu Phe Lys Glu Leu
                30                  35                  40

GTT ACC CAG CAG TTG CCC CAT CTG CTC AAG GAT GTG GGC TCT CTT GAT          315
Val Thr Gln Gln Leu Pro His Leu Leu Lys Asp Val Gly Ser Leu Asp
            45                  50                  55

GAG AAG ATG AAG AGC TTG GAT GTG AAT CAG GAC TCG GAG CTC AAG TTC          363
Glu Lys Met Lys Ser Leu Asp Val Asn Gln Asp Ser Glu Leu Lys Phe
        60                  65                  70

AAT GAG TAC TGG AGA TTG ATT GGG GAG CTG GCC AAG GAA ATC AGG AAG          411
Asn Glu Tyr Trp Arg Leu Ile Gly Glu Leu Ala Lys Glu Ile Arg Lys
    75                  80                  85

AAG AAA GAC CTG AAG ATC AGG AAG AAG TAAAGCCGCC TGGCTGAGAT               458
Lys Lys Asp Leu Lys Ile Arg Lys Lys
 90                  95

GGGGTGGGCA GGGCAGAGGT GATTCAGGGC CGAGCAGAAC CGCACTCTTT CCCAAATAAA       518

GTTTCCTCCT TGAA                                                        532
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 98 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Ala Ala Glu Pro Leu Thr Glu Leu Glu Glu Ser Ile Glu Thr Val
  1               5                  10                  15

Val Thr Thr Phe Phe Thr Phe Ala Arg Gln Glu Gly Arg Lys Asp Ser
            20                  25                  30

Leu Ser Val Asn Glu Phe Lys Glu Leu Val Thr Gln Gln Leu Pro His
        35                  40                  45

Leu Leu Lys Asp Val Gly Ser Leu Asp Glu Lys Met Lys Ser Leu Asp
    50                  55                  60

Val Asn Gln Asp Ser Glu Leu Lys Phe Asn Glu Tyr Trp Arg Leu Ile
65                  70                  75                  80

Gly Glu Leu Ala Lys Glu Ile Arg Lys Lys Lys Asp Leu Lys Ile Arg
                85                  90                  95

Lys Lys
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CGCCCATGGC AGCAGAACCA CTGA 24

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGCGGATCCC GCAGCAGAAC CACTG 25

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGCGGATCCC GCAGCAGAAC CACTG 25

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CGCGGTACCA GCCAGGCGGC TTTA 24

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGCGGATCCA CCATGGCAGC AGAACCA 27

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGCTCTAGAT CAAGCGTAGT CTGGGACGTC GTATGGGTAC TTCTTCCTGA TCTT 54

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 89 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Pro Ser Glu Leu Glu Lys Ala Leu Ser Asn Leu Ile Asp Val Tyr
1 5 10 15

His Asn Tyr Ser Asn Ile Gln Gly Asn His His Ala Leu Tyr Lys Asn
20 25 30

Asp Phe Lys Lys Met Val Thr Thr Glu Cys Pro Gln Phe Val Gln Asn
35 40 45

Ile Asn Ile Glu Asn Leu Phe Arg Glu Leu Asp Ile Asn Ser Asp Asn
50 55 60

Ala Ile Asn Phe Glu Glu Phe Leu Ala Met Val Ile Lys Val Gly Val
65 70 75 80

Ala Ser His Lys Asp Ser His Lys Glu
85

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CGCAAGCTTA GCCAGGCGGC TTTA 24

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a member selected from the group consisting of:

(a) a polynucleotide having at least a 70% identity to a polynucleotide encoding a polypeptide comprising an amino acid sequence set forth in SEQ ID NO:2;

(b) a polynucleotide which is complementary to the polynucleotide of (a); and

(c) a polynucleotide comprising at least 15 bases of the polynucleotide of (a) or (b).

2. The polynucleotide of Claim 1 wherein the polynucleotide is DNA.

3. The polynucleotide of Claim 1 wherein the polynucleotide is RNA.

4. The polynucleotide of Claim 1 wherein the polynucleotide is genomic DNA.

5. The polynucleotide of Claim 2 which encodes a polypeptide comprising amino acid 1 to 98 of SEQ ID NO:2.

6. An isolated polynucleotide comprising a member selected from the group consisting of:

(a) a polynucleotide which encodes a mature polypeptide having the amino acid sequence expressed by the human cDNA contained in ATCC Deposit No. 97405;

(b) a polynucleotide which is complementary to the polynucleotide of (a); and

(c) a polynucleotide comprising at least 15 bases of the polynucleotide of (a) or (b).

7. The polynucleotide of claim 1 comprising the sequence as set forth in SEQ ID NO:1 from nucleotide 1 to nucleotide 532.

8. The polynucleotide of claim 1 comprising the sequence as set forth in SEQ ID NO:1 from nucleotide 145 to nucleotide 438.

9. A vector comprising the DNA of Claim 2.

10. A host cell comprising the vector of Claim 9.

11. A process for producing a polypeptide comprising: expressing from the host cell of Claim 10 the polypeptide encoded by said DNA.

12. A process for producing a cell which expresses a polypeptide comprising genetically engineering the cell with the vector of Claim 9.

13. A polypeptide comprising a member selected from the group consisting of:

(a) a polypeptide having an amino acid sequence set forth in SEQ ID NO:2; and

(b) a polypeptide which is at least 70% identical to the polypeptide of (a).

14. The polypeptide of Claim 13 wherein the polypeptide comprises amino acid 1 to amino acid 98 of SEQ ID NO:2.

15. A compound which inhibits activation of the polypeptide of claim 13.

16. A compound which activates the polypeptide of claim 13.

17. A method for the treatment of a patient having need of CCII comprising: administering to the patient a therapeutically effective amount of the polypeptide of claim 13.

18. The method of Claim 17 wherein said therapeutically effective amount of the polypeptide is administered by providing to the patient DNA encoding said polypeptide and expressing said polypeptide *in vivo*.

19. A method for the treatment of a patient having need to inhibit a CCII polypeptide comprising: administering to the patient a therapeutically effective amount of the compound of Claim 15.

20. A process for diagnosing a disease or a susceptibility to a disease related to an under-expression of the polypeptide of claim 13 comprising:

determining a mutation in a nucleic acid sequence encoding said polypeptide.

21. A diagnostic process comprising:

analyzing for the presence of the polypeptide of claim 13 in a sample derived from a host.

22. A method for identifying compounds which bind to and inhibit activation of the polypeptide of claim 13 comprising: contacting a cell expressing on the surface thereof a receptor for the polypeptide, said receptor being associated with a second component capable of providing a detectable signal in response to the binding of a compound to said receptor, with an analytically detectable CCII polypeptide and a compound under conditions to permit binding to the receptor; and

determining whether the compound binds to and inhibits the receptor by detecting the absence of a signal generated from the interaction of CCII with the receptor.

FIG. 1

```
          10                  30                  50
CGAGGCAGCTCTCTCCTCCTTCCCCGCTGCTTATAAACCTCAGCCCTGAGGCTCCAGCTC
          70                  90                 110
ACTCTACCCCATCTCCTTGCCGGGTCAGCCCTGACAAAGGTCAGCTAGCCCCTTGAGGAC
         130                 150                 170
ATCAGCTTTGGCCTCAGGGTCCTAATGGCAGCAGAACCACTGACAGAGCTAGAGGAGTCC
                          M   A   A   E   P   L   T   E   L   E   E   S
         190                 210                 230
ATTGAGACCGTGGTCACCACCTTCTTCACCTTTGCAAGGCAGGAGGGCCGGAAGGATAGC
I   E   T   V   V   T   T   F   F   T   F   A   R   Q   E   G   R   K   D   S
         250                 270                 290
CTCAGCGTCAACGAGTTCAAAGAGCTGGTTACCCAGCAGTTGCCCCATCTGCTCAAGGAT
L   S   V   N   E   F   K   E   L   V   T   Q   Q   L   P   H   L   L   K   D
         310                 330                 350

GTGGGCTCTCTTGATGAGAAGATGAAGAGCTTGGATGTGAATCAGGACTCGGAGCTCAAG
V   G   S   L   D   E   K   M   K   S   L   D   V   N   Q   D   S   E   L   K
         370                 390                 410

TTCAATGAGTACTGGAGATTGATTGGGGAGCTGGCCAAGGAAATCAGGAAGAAGAAAGAC
F   N   E   Y   W   R   L   I   G   E   L   A   K   E   I   R   K   K   K   D
         430                 450                 470

CTGAAGATCAGGAAGAAGTAAAGCCGCCTGGCTGAGATGGGGTGGGCAGGGCAGAGGTGA
L   K   I   R   K   K
         490                 510                 530
TTCAGGGCCGAGCAGAACCGCACTCTTTCCCAAATAAAGTTTCCTCCTTGAA
```

FIG. 2

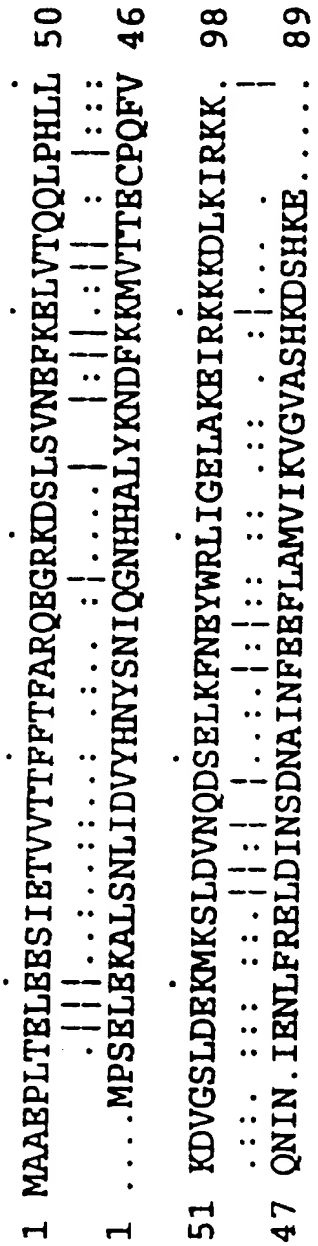
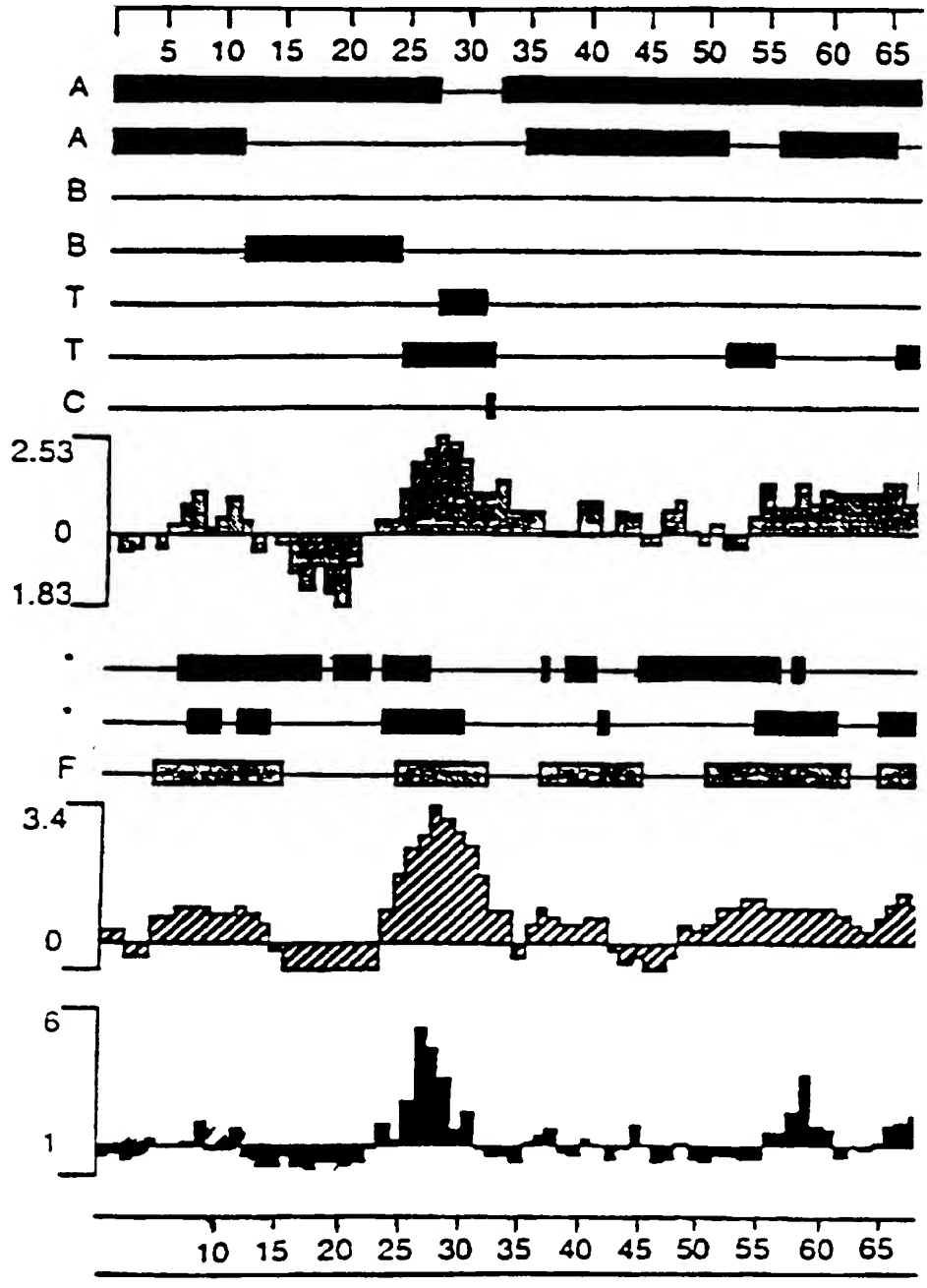
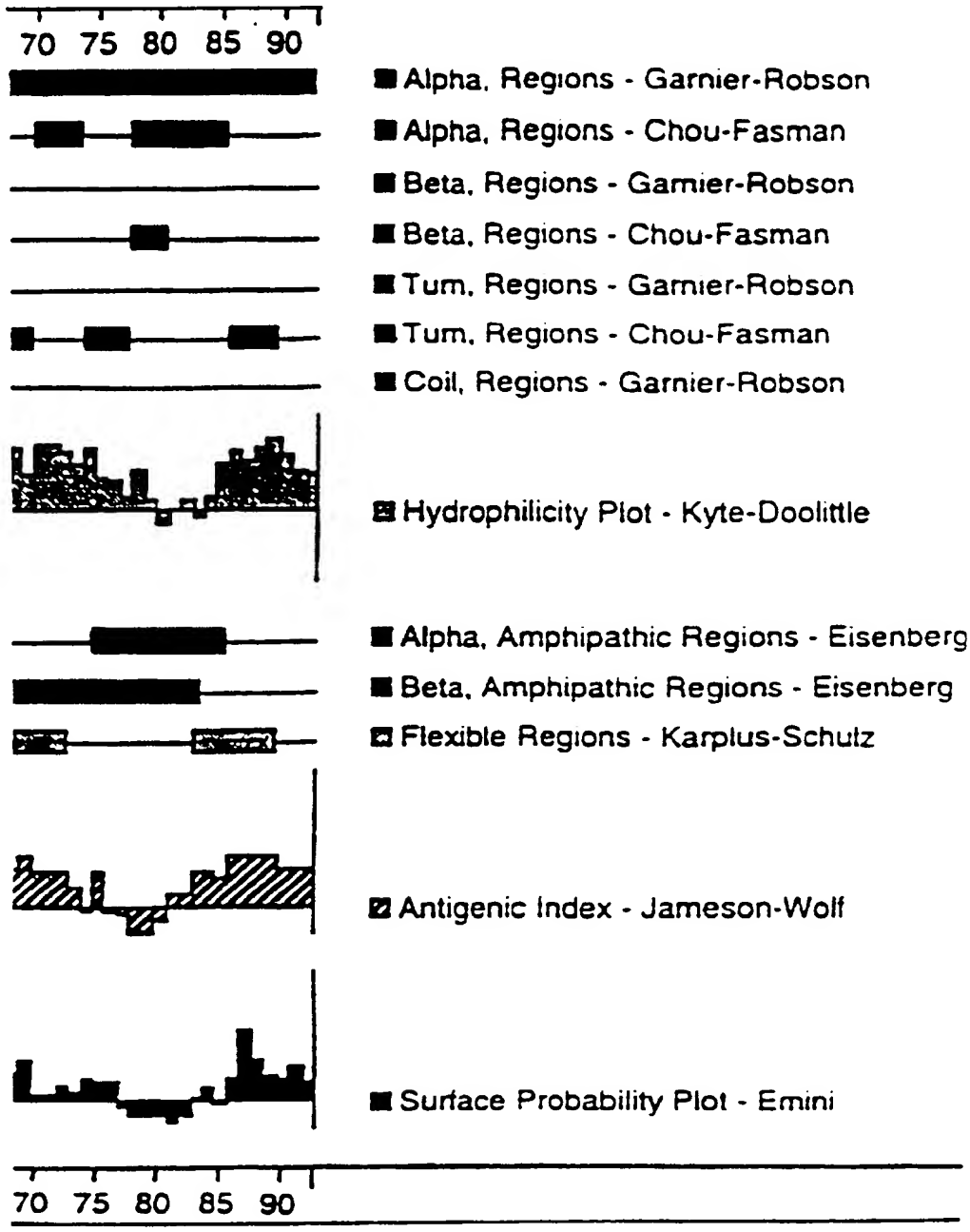


FIG. 3A



MATCH WITH FIG. 3B

FIG. 3B



MATCH WITH FIG. 3A

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/03219

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) Please See Extra Sheet.
US CL Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. 435/69.1, 69.5, 6, 7.1, 7.2, 252.3, 320.1; 536/23.5; 530/351, 350, 389.2; 424/85.1, 514/2, 8, 12, 930/140

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	WO 94/03599 A1 (IWAHORI ET AL) 17 February 1994 (17.02.94), see abstract, figure and claims at page 160, and see sequence alignment.	1-8 -------- 1-14, 17
X --- Y	Database EST-STS on MASPAR search, WashU- Merck EST Project, (St Louis, MO, USA) No. H88181, HILLIER et al. 'Human clone = 252664 5' end, similar to bovine P28783 Calgranulin B,' 21 November 1995, see sequence alignment.	1-8 -------- 1-4, 17
X --- Y	Database EST-STS on MASPAR search, WashU-Merck EST Project (St Louis MO, USA) No. T78482, HILLIER et al. 'cDNA clone 113442 3' end,' 15 March 1995, see sequence alignment.	1-8 -------- 1-14, 17

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 26 JULY 1996	Date of mailing of the international search report 22 AUG 1996
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer GARNETTE D. DRAPER Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/03219

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Database EST-STS on MASPAR search, WashU-Merck EST Project, (St Louis MO, USA), No. R63598, HILLIER et al. 'cDNA clone 138777 5' end,' 26 May 1995, see sequence alignment.	1-14, 17
A, E	US 5,504,003 A (LI ET AL.) 02 April 1996 (02.04.96), see entire document, especially columns 21-22.	1-14, 17
A, E	US 5,525,486 A (HONJO ET AL.) 11 June 1996 (11.06.96), see entire document, especially columns 1-2 and 8.	1-14, 17
A	HARA et al. Molecular cloning and functional characterization of a novel member of the C-C chemokine family. J. Immunol. 01 December 1995, Vol. 155, pages 5352-5358.	1-14, 17
A	PROOST et al. Human and bovine granulocyte chemotactic protein-2: Complete amino acid sequence and fur :ional characterization as chemokines. Biochemistry. 1 -3, Vol. 32, No. 38, pages 10170-10177.	1-14, 17
A	LIAO et al. Human Mig chemokine:Biochemical and functional characterization. J. Exp. Med. November 1995, Vol. 182. No. 5, pages 1301-1314.	1-14, 17
A	SCHULZ-KNAPPE et al. HCC-1, a novel chemokine from human plasma. J. Exp. Med. January 1996, Vol. 183, pages 295-299.	1-14, 17

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/03219

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-14, and 17

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/03219

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (6):

C12P 21/02, 21/06; C12N 1/20, 15/00; C12Q 1/68; G01N 33/53; C07K 1/00, 14/52; A61K 45/05, 38/00, 38/19

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

435/69.1, 69.5, 6, 7.1, 7.2, 252.3, 320.1; 536/23.5; 530/351, 350, 389.2; 424/85.1; 514/2, 8, 12; 930/140

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

APS, CAS ONLINE, DIALOG, MEDLINE, EMBASE, BIOSIS, WPIDS, BIOTECHDS; search terms: chemotactic cytokine or chemokine and (pure, DNA, antibody, antagonist, agonist). Search performed.

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

I. Claims 1-14, and 17, drawn to an isolated polynucleotide, which encodes a chemotactic cytokine, a vector, a host cell, a process for producing chemotactic cytokine polypeptides and cells which express the polypeptide, and uses of the chemotactic cytokine polypeptides.

II. Claims 15 and 19, drawn to a compound which inhibits (antagonist) activation of the cytokine and a method for the treatment of a patient comprising administering the antagonist.

III. Claim 16, drawn to a compound (agonist) which activates the chemotactic cytokine.

IV. Claim 18, drawn to a method for the treatment of a patient comprising administering DNA.

V. Claim 20, drawn to a process for diagnosing a disease comprising determining a mutation in a nucleic acid sequence.

VI. Claim 21, drawn to a diagnostic process comprising analyzing for the chemotactic cytokine polypeptide.

VII. Claim 22, drawn to a method for identifying compounds that are agonist or antagonist of the chemotactic cytokine.

The DNA, the polypeptide and use of the polypeptide compositions of Group I have materially different chemical structures and biological functions from the compound of Group II and compound of Group III. The special features by which the DNA and polypeptide of Group I are defined distinguish them from the special technical features which define the compound of Group II and compound of Group III. The methods of each group, I, II and IV-VII are materially different from the methods of any other groups of the PCT Rule 13.2 so as to form a single inventive concept.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/47, 16/18, A61K 38/17, C12Q 1/68, G01N 33/68

A1

(11) International Publication Number: **WO 98/26068**

(43) International Publication Date: 18 June 1998 (18.06.98)

(21) International Application Number: PCT/US97/22541

(22) International Filing Date: 8 December 1997 (08.12.97)

(30) Priority Data:
08/766,605 12 December 1996 (12.12.96) US

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

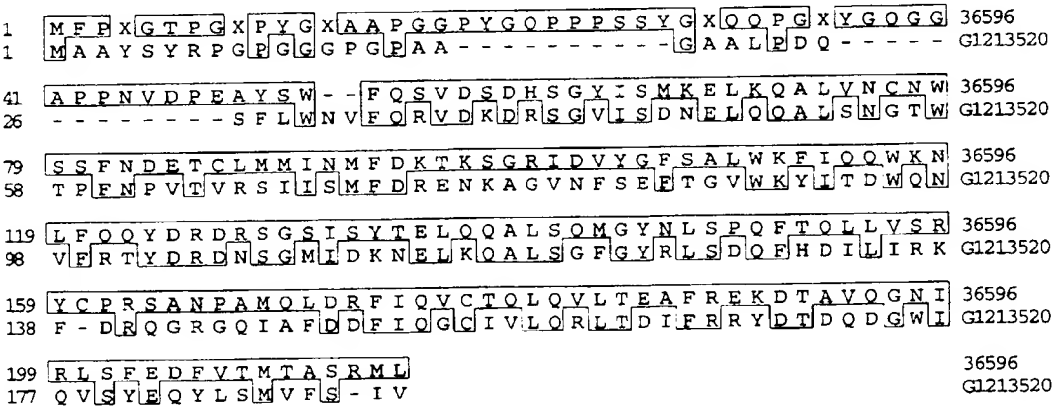
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). GOLI, Surya, K. [IN/US]; 620 Iris Avenue #338, Sunnyvale, CA 94086 (US).

(74) Agent: BILLINGS, Lucy, J.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AT, AU, BR, CA, CH, CN, DE, DK, ES, FI, GB, IL, JP, KR, MX, NO, NZ, RU, SE, SG, US, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: HUMAN APOPTOSIS-RELATED CALCIUM-BINDING PROTEIN



(57) Abstract

The present invention provides a human apoptosis-related calcium-binding protein (HARC) and polynucleotides which identify and encode HARC. The invention also provides genetically engineered expression vectors and host cells comprising the nucleic acid sequences encoding HARC and a method for producing HARC. The invention also provides for agonists, antibodies, or antagonists specifically binding HARC, and their use, in the prevention and treatment of diseases associated with expression of HARC. Additionally, the invention provides for the use of antisense molecules to polynucleotides encoding HARC for the treatment of diseases associated with the expression of HARC. The invention also provides diagnostic assays which utilize the polynucleotide, or fragments or the complement thereof, and antibodies specifically binding HARC.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

HUMAN APOPTOSIS-RELATED CALCIUM-BINDING PROTEIN

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of a novel human apoptosis-related calcium-binding protein and to the use of these sequences in the diagnosis, prevention, and treatment of diseases associated with decreased or increased apoptosis.

BACKGROUND ART

Normal development, growth, and homeostasis in multicellular organisms require a careful balance between the production and destruction of cells in tissues throughout the body. Cell division is a carefully coordinated process with numerous checkpoints and control mechanisms. These mechanisms are designed to regulate DNA replication and to prevent inappropriate or excessive proliferation. In contrast, programmed cell death is the genetically controlled process by which unneeded or damaged cells can be eliminated without causing the tissue destruction and inflammatory responses that are often associated with acute injury and necrosis.

The term "apoptosis" was first used by Kerr, J.F. et al. (1972; Br. J. Cancer 26:239-257) to describe the morphological changes that characterize cells undergoing programmed cell death. Apoptotic cells have a shrunken appearance with an altered membrane lipid content and highly condensed nuclei. Apoptotic cells are rapidly phagocytosed by neighboring cells or macrophages without leaking their potentially damaging contents into the surrounding tissue.

The processes and mechanisms regulating apoptosis are highly conserved throughout the phylogenetic tree. Indeed, much of our current knowledge about apoptosis is derived from studies of the nematode, Caenorhabditis elegans and the fruit fly, Drosophila melanogaster (See for example, Steller, H. (1995) Science 267:1445-1449, and references therein). Dysregulation of apoptosis has recently been recognized as a significant factor in the pathogenesis of human disease. For example, inappropriate cell survival can cause or contribute to many diseases such as cancer, autoimmune diseases, and inflammatory diseases. In contrast, increased apoptosis can cause immunodeficiency diseases such as AIDS, neurodegenerative disorders, and myelodysplastic syndromes (reviewed by Thompson, C.B.

(1995) Science 267:1456-1462).

A variety of ligands and their cellular receptors, enzymes, tumor suppressors, viral gene products, pharmacological agents, and inorganic ions have important positive or negative roles in regulating and implementing the apoptotic destruction of a cell (Steller, H., supra; Thompson, C.B., supra). Although many different extra- and intracellular signals can trigger apoptosis (cf. Raff, M.C. (1992) Nature 356:397-400, Raff, M.C. et al. (1993) Science 262:695-700, and Steller, H. and M.E. Grether, (1994) Neuron 13:1269-1274), these signals probably all converge on a common mechanism that ultimately causes the cell to die.

The mouse apoptosis-linked gene 2 (ALG-2) is one of six clones that has been identified using an in vitro "death trap" model for apoptosis. In this model, 3DO hybridoma cells are transformed with cloned cDNAs and then induced to undergo programmed cell death by cross-linking their T cell receptors. Certain cDNAs afford protection to the transformed 3DO cells either by expressing a protein which inhibits apoptosis or by expressing an antisense RNA which blocks the synthesis of a required protein (Vito, P. et al. (1996) Science 271:521-525).

Analysis of the recovered ALG-2 clone indicates that it is of the latter type, thereby identifying ALG-2 as a protein that is essential for programmed cell death. Northern blot analyses detect a single, ~1.3 KB ALG-2 transcript that is constitutively expressed in all adult mouse tissues; expression is highest in the thymus and liver and lowest in the testes and skeletal muscles. Constitutive expression in normal tissue implies that ALG-2 is probably in an inactive state until the apoptosis pathway is triggered. Apoptosis induced by T cell receptor cross-linking, Fas-Fas ligand interactions, and glucocorticoid treatment all depend on functional ALG-2, which is likely to be a part of the common pathway leading to cell death (Vito, P. et al., supra).

The ALG-2 sequence predicts an acidic protein of 191 amino acids with 2 EF-hand Ca^{2+}-binding domains; both EF-hand domains are required for Ca^{2+} binding. ALG-2 is the first Ca^{2+}-binding protein that has been shown to be directly required in the apoptosis pathway (Vito, P. et al., supra). A requirement for Ca^{2+} in apoptosis was previously suggested by studies showing its involvement in DNA cleavage (Hewish, D.R. and L.A. Burgoyne (1973) Biochem. Biophys. Res. Comm. 52:504-510). Other studies show that: 1) intracellular Ca^{2+}

concentrations increase when apoptosis is triggered in thymocytes by either T cell receptor cross-linking or by glucocorticoids, and 2) cell death can be prevented by blocking this increase in intracellular Ca^{2+} (McConkey, D.J. et al. (1989) J. Immunol. 143:1801-1806; McConkey, D.J. et al. (1989) Arch. Biochem. Biophys. 269:365-370). Additional support for the role of Ca^{2+} in apoptosis comes from work on Fas-mediated cell death (Vignaux, F. et al. (1995) J. Exp. Med. 181:781-786; Oshimi, Y. and S. Miyazaki (1995) J. Immunol. 154:599-609).

The discovery of polynucleotides encoding human apoptosis-related calcium-binding protein, and the molecules themselves, provide a means to investigate the regulation of programmed cell death and apoptosis. Discovery of molecules related to mouse ALG-2 satisfies a need in the art by providing new diagnostic or therapeutic compositions useful in the detection, prevention, and treatment of cancer, autoimmune diseases, lymphoproliferative disorders, psoriasis, atherosclerosis, restenosis, AIDS, immunodeficiency diseases, ischemic injuries, neurodegenerative diseases, osteoporosis, myelodysplastic syndromes, toxin-induced diseases, cachexia, and viral infections.

DISCLOSURE OF THE INVENTION

The present invention features a novel human apoptosis-related calcium-binding protein hereinafter designated HARC and characterized as having similarity to mouse apoptosis-linked gene 2, ALG-2.

Accordingly, the invention features a substantially purified HARC having the amino acid sequence shown in SEQ ID NO:1.

One aspect of the invention features isolated and substantially purified polynucleotides that encode HARC. In a particular aspect, the polynucleotide is the nucleotide sequence of SEQ ID NO:2.

The invention also relates to a polynucleotide sequence comprising the complement of SEQ ID NO:2 or variants thereof. In addition, the invention features polynucleotide sequences which hybridize under stringent conditions to SEQ ID NO:2.

The invention additionally features nucleic acid sequences encoding polypeptides, oligonucleotides, peptide nucleic acids (PNA), fragments, portions or antisense molecules thereof, and expression vectors and host cells comprising polynucleotides that encode HARC.

The present invention also features antibodies which bind specifically to HARC, and pharmaceutical compositions comprising substantially purified HARC. The invention also features the use of agonists and antagonists of HARC.

BRIEF DESCRIPTION OF DRAWINGS

Figures 1A and 1B shows the amino acid sequence (SEQ ID NO:1) and nucleic acid sequence (SEQ ID NO:2) of HARC. The alignment was produced using MacDNASIS PRO™ software (Hitachi Software Engineering Co., Ltd., San Bruno, CA).

Figure 2 shows the amino acid sequence alignments between HARC (SEQ ID NO:1) and mouse ALG-2 (GI1213520; SEQ ID NO:3). The alignment was produced using the multisequence alignment program of DNASTAR™ software (DNASTAR Inc, Madison WI).

Figure 3 shows the hydrophobicity plot (MacDNASIS PRO software) for HARC, SEQ ID NO: 1; the positive X axis reflects amino acid position, and the negative Y axis, hydrophobicity.

Figure 4 shows the hydrophobicity plot for ALG-2, SEQ ID NO:3.

Figures 5A and 5B shows the northern analysis for SEQ ID NO:2. The northern analysis was produced electronically using the LIFESEQ™ database (Incyte Pharmaceuticals, Inc., Palo Alto, CA).

MODES FOR CARRYING OUT THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular methodology, protocols, cell lines, vectors, and reagents described as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a", "an", and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a host cell" includes a plurality of such host cells, reference to the "antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same

meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods, devices, and materials are now described. All publications mentioned herein are incorporated herein by reference for the purpose of describing and disclosing the cell lines, vectors, and methodologies which are reported in the publications which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"Nucleic acid sequence" as used herein refers to an oligonucleotide, nucleotide, or polynucleotide, and fragments or portions thereof, and to DNA or RNA of genomic or synthetic origin which may be single- or double-stranded, and represent the sense or antisense strand. Similarly, "amino acid sequence" as used herein refers to an oligopeptide, peptide, polypeptide, or protein sequence, and fragments or portions thereof, and to naturally occurring or synthetic molecules.

Where "amino acid sequence" is recited herein to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms, such as "polypeptide" or "protein" are not meant to limit the amino acid sequence to the complete, native amino acid sequence associated with the recited protein molecule.

"Peptide nucleic acid", as used herein, refers to a molecule which comprises an oligomer to which an amino acid residue, such as lysine, and an amino group have been added. These small molecules, also designated anti-gene agents, stop transcript elongation by binding to their complementary strand of nucleic acid (Nielsen, P.E. et al. (1993) Anticancer Drug Des. 8:53-63).

HARC, as used herein, refers to the amino acid sequences of substantially purified HARC obtained from any species, particularly mammalian, including bovine, ovine, porcine, murine, equine, and preferably human, from any source whether natural, synthetic, semi-synthetic, or recombinant.

"Consensus", as used herein, refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, or which has been extended using XL-PCR™ (Perkin

Elmer, Norwalk, CT) in the 5′ and/or the 3′ direction and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte clone using the GELVIEW™ Fragment Assembly system (GCG, Madison, WI), or which has been both extended and assembled.

A "variant" of HARC, as used herein, refers to an amino acid sequence that is altered by one or more amino acids. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties, e.g., replacement of leucine with isoleucine. More rarely, a variant may have "nonconservative" changes, e.g., replacement of a glycine with a tryptophan. Similar minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, DNASTAR software.

A "deletion", as used herein, refers to a change in either amino acid or nucleotide sequence in which one or more amino acid or nucleotide residues, respectively, are absent.

An "insertion" or "addition", as used herein, refers to a change in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid or nucleotide residues, respectively, as compared to the naturally occurring molecule.

A "substitution", as used herein, refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

The term "biologically active", as used herein, refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic HARC, or any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "agonist", as used herein, refers to a molecule which, when bound to HARC, causes a change in HARC which modulates the activity of HARC. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to HARC.

The terms "antagonist" or "inhibitor", as used herein, refer to a molecule which, when bound to HARC, blocks or modulates the biological or immunological activity of HARC.

Antagonists and inhibitors may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to HARC.

The term "modulate", as used herein, refers to a change or an alteration in the biological activity of HARC. Modulation may be an increase or a decrease in protein activity, a change in binding characteristics, or any other change in the biological, functional or immunological properties of HARC.

The term "mimetic", as used herein, refers to a molecule, the structure of which is developed from knowledge of the structure of HARC or portions thereof and, as such, is able to effect some or all of the actions of ALG-2-like molecules.

The term "derivative", as used herein, refers to the chemical modification of a nucleic acid encoding HARC or the encoded HARC. Illustrative of such modifications would be replacement of hydrogen by an alkyl, acyl, or amino group. A nucleic acid derivative would encode a polypeptide which retains essential biological characteristics of the natural molecule.

The term "substantially purified", as used herein, refers to nucleic or amino acid sequences that are removed from their natural environment, isolated or separated, and are at least 60% free, preferably 75% free, and most preferably 90% free from other components with which they are naturally associated.

"Amplification" as used herein refers to the production of additional copies of a nucleic acid sequence and is generally carried out using polymerase chain reaction (PCR) technologies well known in the art (Dieffenbach, C.W. and G.S. Dveksler (1995) PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview, NY).

The term "hybridization", as used herein, refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

The term "hybridization complex", as used herein, refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen binds between complementary G and C bases and between complementary A and T bases; these hydrogen bonds may be further stabilized by base stacking interactions. The two complementary nucleic acid sequences hydrogen bond in an antiparallel configuration. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or between one nucleic acid

sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., membranes, filters, chips, pins or glass slides to which cells have been fixed for in situ hybridization).

The terms "complementary" or "complementarity", as used herein, refer to the natural binding of polynucleotides under permissive salt and temperature conditions by base-pairing. For example, for the sequence "A-G-T" binds to the complementary sequence "T-C-A". Complementarity between two single-stranded molecules may be "partial", in which only some of the nucleic acids bind, or it may be complete when total complementarity exists between the single-stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands.

The term "homology", as used herein, refers to a degree of complementarity. There may be partial homology or complete homology (i.e., identity). A partially complementary sequence is one that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid; it is referred to using the functional term "substantially homologous." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization and the like) under conditions of low stringency. A substantially homologous sequence or probe will compete for and inhibit the binding (i.e., the hybridization) of a completely homologous sequence or probe to the target sequence under conditions of low stringency. This is not to say that conditions of low stringency are such that non-specific binding is permitted; low stringency conditions require that the binding of two sequences to one another be a specific (i.e., selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% identity); in the absence of non-specific binding, the probe will not hybridize to the second non-complementary target sequence.

As known in the art, numerous equivalent conditions may be employed to comprise either low or high stringency conditions. Factors such as the length and nature (DNA, RNA, base composition) of the sequence, nature of the target (DNA, RNA, base composition,

presence in solution or immobilization, etc.), and the concentration of the salts and other components (e.g., the presence or absence of formamide, dextran sulfate and/or polyethylene glycol) are considered and the hybridization solution may be varied to generate conditions of either low or high stringency different from, but equivalent to, the above listed conditions.

The term "stringent conditions", as used herein, is the "stringency" which occurs within a range from about Tm-5°C (5°C below the melting temperature (Tm) of the probe) to about 20°C to 25°C below Tm. As will be understood by those of skill in the art, the stringency of hybridization may be altered in order to identify or detect identical or related polynucleotide sequences.

The term "antisense", as used herein, refers to nucleotide sequences which are complementary to a specific DNA or RNA sequence. The term "antisense strand" is used in reference to a nucleic acid strand that is complementary to the "sense" strand. Antisense molecules may be produced by any method, including synthesis by ligating the gene(s) of interest in a reverse orientation to a viral promoter which permits the synthesis of a complementary strand. Once introduced into a cell, this transcribed strand combines with natural sequences produced by the cell to form duplexes. These duplexes then block either the further transcription or translation. In this manner, mutant phenotypes may be generated. The designation "negative" is sometimes used in reference to the antisense strand, and "positive" is sometimes used in reference to the sense strand.

The term "portion", as used herein, with regard to a protein (as in "a portion of a given protein") refers to fragments of that protein. The fragments may range in size from four amino acid residues to the entire amino acid sequence minus one amino acid. Thus, a protein "comprising at least a portion of the amino acid sequence of SEQ ID NO:1" encompasses the full-length human HARC and fragments thereof.

"Transformation", as defined herein, describes a process by which exogenous DNA enters and changes a recipient cell. It may occur under natural or artificial conditions using various methods well known in the art. Transformation may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method is selected based on the host cell being transformed and may include, but is not limited to, viral infection, electroporation, lipofection, and particle bombardment. Such

"transformed" cells include stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome. They also include cells which transiently express the inserted DNA or RNA for limited periods of time.

The term "antigenic determinant", as used herein, refers to that portion of a molecule that makes contact with a particular antibody (i.e., an epitope). When a protein or fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to a given region or three-dimensional structure on the protein; these regions or structures are referred to as antigenic determinants. An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The terms "specific binding" or "specifically binding", as used herein, in reference to the interaction of an antibody and a protein or peptide, mean that the interaction is dependent upon the presence of a particular structure (i.e., the antigenic determinant or epitope) on the protein; in other words, the antibody is recognizing and binding to a specific protein structure rather than to proteins in general. For example, if an antibody is specific for epitope "A", the presence of a protein containing epitope A (or free, unlabeled A) in a reaction containing labeled "A" and the antibody will reduce the amount of labeled A bound to the antibody.

The term "sample", as used herein, is used in its broadest sense. A biological sample suspected of containing nucleic acid encoding HARC or fragments thereof may comprise a cell, chromosomes isolated from a cell (e.g., a spread of metaphase chromosomes), genomic DNA (in solution or bound to a solid support such as for Southern analysis), RNA (in solution or bound to a solid support such as for northern analysis), cDNA (in solution or bound to a solid support), an extract from cells or a tissue, and the like.

The term "correlates with expression of a polynucleotide", as used herein, indicates that the detection of the presence of ribonucleic acid that is similar to SEQ ID NO:2 by northern analysis is indicative of the presence of mRNA encoding HARC in a sample and thereby correlates with expression of the transcript from the polynucleotide encoding the protein.

"Alterations" in the polynucleotide of SEQ ID NO: 2, as used herein, comprise any

alteration in the sequence of polynucleotides encoding HARC including deletions, insertions, and point mutations that may be detected using hybridization assays. Included within this definition is the detection of alterations to the genomic DNA sequence which encodes HARC (e.g., by alterations in the pattern of restriction fragment length polymorphisms capable of hybridizing to SEQ ID NO:2), the inability of a selected fragment of SEQ ID NO: 2 to hybridize to a sample of genomic DNA (e.g., using allele-specific oligonucleotide probes), and improper or unexpected hybridization, such as hybridization to a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HARC (e.g., using fluorescent in situ hybridization (FISH) to metaphase chromosomes spreads).

As used herein, the term "antibody" refers to intact molecules as well as fragments thereof, such as Fa, $F(ab')_2$, and Fv, which are capable of binding the epitopic determinant. Antibodies that bind HARC polypeptides can be prepared using intact polypeptides or fragments containing small peptides of interest as the immunizing antigen. The polypeptide or peptide used to immunize an animal can be derived from the transition of RNA or synthesized chemically, and can be conjugated to a carrier protein, if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin and thyroglobulin. The coupled peptide is then used to immunize the animal (e.g., a mouse, a rat, or a rabbit).

The term "humanized antibody", as used herein, refers to antibody molecules in which amino acids have been replaced in the non-antigen binding regions in order to more closely resemble a human antibody, while still retaining the original binding ability.

THE INVENTION

The invention is based on the discovery of a novel human apoptosis-related calcium-binding protein (HARC), the polynucleotides encoding HARC, and the use of these compositions for the diagnosis, prevention, and treatment of cancer, autoimmune diseases, lymphoproliferative disorders, psoriasis, atherosclerosis, restenosis, AIDS, immunodeficiency diseases, ischemic injuries, neurodegenerative diseases, osteoporosis, myelodysplastic syndromes, toxin-induced diseases, cachexia, and viral infections.

Nucleic acids encoding the human HARC of the present invention were first identified in Incyte Clone 036596 from the human umbilical cord vein endothelial cell line cDNA

library (HUVENOB01) through a computer-generated search for amino acid sequence alignments. A consensus sequence, SEQ ID NO:2, was derived from the following overlapping and/or extended nucleic acid sequences: Incyte Clones 036596 (HUVENOB01), 179146 (PLACNOB01), 334277 (EOSIHET02), 1304460 (PLACNOT02), and 1752846 (LIVRTUT01).

In one embodiment, the invention encompasses a polypeptide comprising the amino acid sequence of SEQ ID NO:1, as shown in Fig. 1. HARC is 214 amino acids in length and has two potential EF-hand calcium-binding domains at residues 53-69 and 117-137. HARC has chemical and structural homology with ALG-2 (GI1213520; SEQ ID NO:3). In particular, HARC and ALG-2 share 47% overall identity, and 58% identity in each of the two calcium-binding domains. As illustrated by Figs. 3 and 4, HARC and ALG-2 have similar hydrophobicity plots, and both are very acidic proteins with isoelectric points of 5.26 and 5.01, respectively. Northern analysis (Fig. 5) shows the expression of this sequence in various libraries, 55% of which are from cancers, tissues associated with inflammatory diseases, fetal tissues, and cell lines. Of particular note is the expression of HARC in normal brain (5/44) and prostate (3/44) cDNA libraries.

The invention also encompasses HARC variants. A preferred HARC variant is one having at least 80%, and more preferably 90%, amino acid sequence similarity to the HARC amino acid sequence (SEQ ID NO:1). A most preferred HARC variant is one having at least 95% amino acid sequence similarity to SEQ ID NO:1.

The invention also encompasses polynucleotides which encode HARC. Accordingly, any nucleic acid sequence which encodes the amino acid sequence of HARC can be used to generate recombinant molecules which express HARC. In a particular embodiment, the invention encompasses the polynucleotide comprising the nucleic acid sequence of SEQ ID NO:2 as shown in Figures 1A and 1B.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of nucleotide sequences encoding HARC, some bearing minimal homology to the nucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of nucleotide sequence that could be made by selecting combinations based on possible codon choices.

These combinations are made in accordance with the standard triplet genetic code as applied to the nucleotide sequence of naturally occurring HARC, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode HARC and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring HARC under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HARC or its derivatives possessing a substantially different codon usage. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HARC and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences, or portions thereof, which encode HARC and its derivatives, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents that are well known in the art at the time of the filing of this application. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HARC or any portion thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed nucleotide sequences, and in particular, those shown in SEQ ID NO:2, under various conditions of stringency. Hybridization conditions are based on the melting temperature (Tm) of the nucleic acid binding complex or probe, as taught in Wahl, G.M. and S.L. Berger (1987; Methods Enzymol. 152:399-407) and Kimmel, A.R. (1987; Methods Enzymol. 152:507-511), and may be used at a defined stringency.

Altered nucleic acid sequences encoding HARC which are encompassed by the invention include deletions, insertions, or substitutions of different nucleotides resulting in a polynucleotide that encodes the same or a functionally equivalent HARC. The encoded protein may also contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HARC. Deliberate amino acid

substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues as long as the biological activity of HARC is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid; positively charged amino acids may include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; phenylalanine and tyrosine.

Also included within the scope of the present invention are alleles of the genes encoding HARC. As used herein, an "allele" or "allelic sequence" is an alternative form of the gene which may result from at least one mutation in the nucleic acid sequence. Alleles may result in altered mRNAs or polypeptides whose structure or function may or may not be altered. Any given gene may have none, one, or many allelic forms. Common mutational changes which give rise to alleles are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

Methods for DNA sequencing which are well known and generally available in the art may be used to practice any embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, Sequenase® (US Biochemical Corp, Cleveland, OH), Taq polymerase (Perkin Elmer), thermostable T7 polymerase (Amersham, Chicago, IL), or combinations of recombinant polymerases and proofreading exonucleases such as the ELONGASE Amplification System marketed by Gibco BRL (Gaithersburg, MD). Preferably, the process is automated with machines such as the Hamilton Micro Lab 2200 (Hamilton, Reno, NV), Peltier Thermal Cycler (PTC200; M.J. Research, Watertown, MA) and the ABI 377 DNA sequencers (Perkin Elmer).

The nucleic acid sequences encoding HARC may be extended utilizing a partial nucleotide sequence and employing various methods known in the art to detect upstream sequences such as promoters and regulatory elements. For example, one method which may be employed, "restriction-site" PCR, uses universal primers to retrieve unknown sequence adjacent to a known locus (Sarkar, G. (1993) PCR Methods Applic. 2:318-322). In particular, genomic DNA is first amplified in the presence of primer to linker sequence and a primer

specific to the known region. The amplified sequences are then subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR may also be used to amplify or extend sequences using divergent primers based on a known region (Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186). The primers may be designed using OLIGO 4.06 Primer Analysis software (National Biosciences Inc., Plymouth, MN), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68°-72° C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Another method which may be used is capture PCR which involves PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA (Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119). In this method, multiple restriction enzyme digestions and ligations may also be used to place an engineered double-stranded sequence into an unknown portion of the DNA molecule before performing PCR.

Another method which may be used to retrieve unknown sequences is that of Parker, J.D. et al. (1991; Nuc. Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PromoterFinder™ libraries to walk in genomic DNA (Clontech, Palo Alto, CA). This process avoids the need to screen libraries and is useful in finding intron/exon junctions.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. Also, random-primed libraries are preferable, in that they will contain more sequences which contain the 5′ regions of genes. Use of a randomly primed library may be especially preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into the 5′ and 3′ non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In

particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled devise camera. Output/light intensity may be converted to electrical signal using appropriate software (e.g. Genotyper™ and Sequence Navigator™, Perkin Elmer) and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HARC, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct expression of HARC in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express HARC.

As will be understood by those of skill in the art, it may be advantageous to produce HARC-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter HARC encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding HARC may be ligated to a heterologous sequence to encode a fusion

protein. For example, to screen peptide libraries for inhibitors of HARC activity, it may be useful to encode a chimeric HARC protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the HARC encoding sequence and the heterologous protein sequence, so that HARC may be cleaved and purified away from the heterologous moiety.

In another embodiment, sequences encoding HARC may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M.H. et al. (1980) Nuc. Acids Res. Symp. Ser. 215-223, Horn, T. et al. (1980) Nuc. Acids Res. Symp. Ser. 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of HARC, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J.Y. et al. (1995) Science 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer).

The newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (e.g., Creighton, T. (1983) Proteins, Structures and Molecular Principles, WH Freeman and Co., New York, NY). The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (e.g., the Edman degradation procedure; Creighton, supra). Additionally, the amino acid sequence of HARC, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a biologically active HARC, the nucleotide sequences encoding HARC or functional equivalents, may be inserted into appropriate expression vector, i.e., a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence.

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding HARC and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor

Press, Plainview, NY, and Ausubel, F.M. et al. (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY.

A variety of expression vector/host systems may be utilized to contain and express sequences encoding HARC. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (e.g., baculovirus); plant cell systems transformed with virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems.

The "control elements" or "regulatory sequences" are those non-translated regions of the vector--enhancers, promoters, 5′ and 3′ untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the Bluescript® phagemid (Stratagene, La Jolla, CA) or pSport1™ plasmid (Gibco BRL) and the like may be used. The baculovirus polyhedrin promoter may be used in insect cells. Promoters or enhancers derived from the genomes of plant cells (e.g., heat shock, RUBISCO; and storage protein genes) or from plant viruses (e.g., viral promoters or leader sequences) may be cloned into the vector. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are preferable. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding HARC, vectors based on SV40 or EBV may be used with an appropriate selectable marker.

In bacterial systems, a number of expression vectors may be selected depending upon the use intended for HARC. For example, when large quantities of HARC are needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional E. coli cloning and expression vectors such as Bluescript® (Stratagene), in which the sequence encoding HARC may be ligated into the vector in frame with sequences for the

amino-terminal Met and the subsequent 7 residues of ß-galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509); and the like. pGEX vectors (Promega, Madison, WI) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, Saccharomyces cerevisiae, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel et al. (supra) and Grant et al. (1987) Methods Enzymol. 153:516-544.

In cases where plant expression vectors are used, the expression of sequences encoding HARC may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or Murry, L.E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, NY; pp. 191-196.)

An insect system may also be used to express HARC. For example, in one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in Spodoptera frugiperda cells or in Trichoplusia larvae. The sequences encoding HARC may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of HARC will render the polyhedrin gene inactive and produce recombinant virus lacking

coat protein. The recombinant viruses may then be used to infect, for example, S. frugiperda cells or Trichoplusia larvae in which HARC may be expressed (Engelhard, E.K. et al. (1994) Proc. Nat. Acad. Sci. 91:3224-3227).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HARC may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing HARC in infected host cells (Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HARC. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding HARC, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, HEK293, and WI38, which have specific cellular machinery and characteristic mechanisms for such

post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express HARC may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. et al. (1977) Cell 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. et al. (1980) Cell 22:817-23) genes which can be employed in tk^- or $aprt^-$ cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. 77:3567-70); npt, which confers resistance to the aminoglycosides neomycin and G-418 (Colbere-Garapin, F. et al (1981) J. Mol. Biol. 150:1-14) and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, supra). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-51). Recently, the use of visible markers has gained popularity with such markers as anthocyanins, ß glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C.A. et al. (1995) Methods Mol. Biol. 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of

interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding HARC is inserted within a marker gene sequence, recombinant cells containing sequences encoding HARC can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding HARC under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells which contain the nucleic acid sequence encoding HARC and express HARC may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

The presence of polynucleotide sequences encoding HARC can be detected by DNA-DNA or DNA-RNA hybridization or amplification using probes or portions or fragments of polynucleotides encoding HARC. Nucleic acid amplification based assays involve the use of oligonucleotides or oligomers based on the sequences encoding HARC to detect transformants containing DNA or RNA encoding HARC. As used herein "oligonucleotides" or "oligomers" refer to a nucleic acid sequence of at least about 10 nucleotides and as many as about 60 nucleotides, preferably about 15 to 30 nucleotides, and more preferably about 20-25 nucleotides, which can be used as a probe or amplimer.

A variety of protocols for detecting and measuring the expression of HARC, using either polyclonal or monoclonal antibodies specific for the protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HARC is preferred, but a competitive binding assay may be employed. These and other assays are described, among other places, in Hampton, R. et al. (1990; Serological Methods, a Laboratory Manual, APS Press, St Paul, MN) and Maddox, D.E. et al. (1983; J. Exp. Med. 158:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing

labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HARC include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding HARC, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits (Pharmacia & Upjohn, (Kalamazoo, MI); Promega (Madison WI); and U.S. Biochemical Corp., Cleveland, OH). Suitable reporter molecules or labels, which may be used, include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding HARC may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HARC may be designed to contain signal sequences which direct secretion of HARC through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding HARC to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, WA). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, CA) between the purification domain and HARC may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing HARC and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography as described in Porath, J. et al. (1992, Prot. Exp. Purif. 3: 263-281) while the

enterokinase cleavage site provides a means for purifying HARC from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D.J. et al. (1993; DNA Cell Biol. 12:441-453).

In addition to recombinant production, fragments of HARC may be produced by direct peptide synthesis using solid-phase techniques Merrifield J. (1963) J. Am. Chem. Soc. 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Various fragments of HARC may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

THERAPEUTICS

Based on the chemical and structural homology between HARC (SEQ ID NO:1) and mouse ALG-2 (SEQ ID NO:3), HARC appears to play a role in the development of diseases related to decreased or increased apoptosis. Diseases arising from decreased apoptosis include, but are not limited to, cancer of the breast, prostate, and ovary as well as lymphomas and carcinomas, autoimmune disorders such as systemic lupus erythematosus, glomerulonephritis, Sjögren's syndrome, Graves disease, multiple sclerosis, rheumatoid arthritis, and diabetes, inflammatory diseases such as osteoarthritis, Crohn's disease, inflammatory bowel disease, and colitis, proliferative disorders such as atherosclerosis, restenosis, psoriasis, lymphadenopathy, and viral infections such as by herpesviruses, poxviruses, and adenoviruses.

HARC-associated diseases that are related to increased apoptosis include, but are not limited to, AIDS and other infectious or genetic immunodeficiencies, neurodegenerative diseases such as Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, retinitis pigmentosa, and cerebellar degeneration, myelodysplastic syndromes such as aplastic anemia, ischemic injuries such as myocardial infarction, stroke, and reperfusion injury, toxin-induced diseases such as alcohol-induced liver damage, cirrhosis, and lathyrism, wasting diseases such as cachexia, viral infections such as by hepatitis B and C, and osteoporosis.

Therefore, in one embodiment, HARC or a fragment or derivative thereof may be administered to a subject to treat the disorders and diseases resulting from decreased

apoptosis including, but not limited to, those listed above.

In another embodiment, a vector capable of expressing HARC, or a fragment or a derivative thereof, may also be administered to a subject to treat the disorders and diseases resulting from decreased apoptosis including, but not limited to, those described above.

In another embodiment, agonists of HARC may be administered to a subject to treat or prevent the disorders and diseases resulting from decreased apoptosis including, but not limited to, those described above.

In another embodiment, antagonists or inhibitors of HARC may be administered to a subject to treat or prevent the disorders and diseases resulting from increased apoptosis including, but not limited to, those described above.

In another embodiment, a vector expressing antisense of the polynucleotide encoding HARC may be administered to a subject to treat or prevent the disorders and diseases resulting from increased apoptosis including, but not limited to, those described above.

In one aspect, antibodies which are specific for HARC may be used directly as an antagonist. In another aspect, antibodies may be used to deliver a pharmaceutical agent to cells or tissue which express HARC.

In other embodiments, any of the therapuetic compositions described above may be administered in combination with other conventional therapeutic agents. The combination of therapeutic agents having different mechanisms of action will have synergistic effects allowing for the use of lower effective doses of each agent and thereby lessening side effects.

Antagonists or inhibitors of HARC may be produced using methods which are generally known in the art. In particular, purified HARC may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HARC.

Antibodies which are specific for HARC may be used directly as an antagonist, or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express HARC. The antibodies may be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies, (i.e., those which inhibit dimer formation) are

especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others, may be immunized by injection with HARC or any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the peptides, fragments, or oligopeptides used to induce antibodies to HARC have an amino acid sequence consisting of at least five amino acids, and more preferably at least 10 amino acids. It is also preferable that they are identical to a portion of the amino acid sequence of the natural protein, and they may contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of HARC amino acids may be fused with those of another protein such as keyhole limpet hemocyanin and antibody produced against the chimeric molecule.

Monoclonal antibodies to HARC may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique (Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. 80:2026-2030; Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120).

In addition, techniques developed for the production of "chimeric antibodies", the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity can be used (Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; Takeda, S. et al. (1985) Nature 314:452-454). Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce HARC-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from

random combinatorial immunoglobin libraries (Burton D.R. (1991) Proc. Natl. Acad. Sci. 88:11120-3).

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening recombinant immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature (Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. 86: 3833-3837; Winter, G. et al. (1991) Nature 349:293-299).

Antibody fragments which contain specific binding sites for HARC may also be generated. For example, such fragments include, but are not limited to, the F(ab')2 fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse, W.D. et al. (1989) Science 254:1275-1281).

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HARC and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HARC epitopes is preferred, but a competitive binding assay may also be employed (Maddox, supra).

In another embodiment of the invention, the polynucleotides encoding HARC, or any fragment thereof, or antisense molecules, may be used for therapeutic purposes. In one aspect, antisense to the polynucleotide encoding HARC may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HARC. Thus, antisense molecules may be used to modulate HARC activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligomers or larger fragments, can be designed from various locations along the coding or control regions of sequences encoding HARC.

Expression vectors derived from retroviruses, adenovirus, herpes or vaccinia viruses,

or from various bacterial plasmids may be used for delivery of nucleotide sequences to the targeted organ, tissue or cell population. Methods which are well known to those skilled in the art can be used to construct recombinant vectors which will express antisense molecules complementary to the polynucleotides of the gene encoding HARC. These techniques are described both in Sambrook et al. (supra) and in Ausubel et al. (supra).

Genes encoding HARC can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide or fragment thereof which encodes HARC. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector and even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing antisense molecules, DNA, RNA, or PNA, to the control regions of the gene encoding HARC, i.e., the promoters, enhancers, and introns. Oligonucleotides derived from the transcription initiation site, e.g., between positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using "triple helix" base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature (Gee, J.E. et al. (1994) In: Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY). The antisense molecules may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. Examples which may be used include engineered hammerhead motif ribozyme molecules that can specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HARC.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites which include the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Antisense molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HARC. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize antisense RNA constitutively or inducibly can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5′ and/or 3′ ends of the molecule or the use of phosphorothioate or 2′-O-methyl phosphodiester linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection and by liposome injections may be achieved using methods which are well known in the art.

Any of the therapeutic methods described above may be applied to any subject in need

of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HARC, antibodies to HARC, mimetics, agonists, antagonists, or inhibitors of HARC. The compositions may be administered alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipients, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars,

including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks'solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a

manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1-50 mM histidine, 0.1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HARC, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models, usually mice, rabbits, dogs, or pigs. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HARC or fragments thereof, antibodies of HARC, agonists, antagonists or inhibitors of HARC, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50.

Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range

of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind HARC may be used for the diagnosis of conditions or diseases characterized by expression of HARC, or in assays to monitor patients being treated with HARC, agonists, antagonists or inhibitors. The antibodies useful for diagnostic purposes may be prepared in the same manner as those described above for therapeutics. Diagnostic assays for HARC include methods which utilize the antibody and a label to detect HARC in human body fluids or extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by joining them, either covalently or non-covalently, with a reporter molecule. A wide variety of reporter molecules which are known in the art may be used, several of which are described above.

A variety of protocols including ELISA, RIA, and FACS for measuring HARC are known in the art and provide a basis for diagnosing altered or abnormal levels of HARC

expression. Normal or standard values for HARC expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to HARC under conditions suitable for complex formation The amount of standard complex formation may be quantified by various methods, but preferably by photometric, means. Quantities of HARC expressed in subject, control and disease, samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding HARC may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, antisense RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of HARC may be correlated with disease. The diagnostic assay may be used to distinguish between absence, presence, and excess expression of HARC, and to monitor regulation of HARC levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HARC or closely related molecules, may be used to identify nucleic acid sequences which encode HARC. The specificity of the probe, whether it is made from a highly specific region, e.g., 10 unique nucleotides in the 5′ regulatory region, or a less specific region, e.g., especially in the 3′ coding region, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low) will determine whether the probe identifies only naturally occurring sequences encoding HARC, alleles, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably contain at least 50% of the nucleotides from any of the HARC encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and derived from the nucleotide sequence of SEQ ID NO:2 or from genomic sequence including promoter, enhancer elements, and introns of the naturally occurring HARC.

Means for producing specific hybridization probes for DNAs encoding HARC include the cloning of nucleic acid sequences encoding HARC or HARC derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, commercially available,

and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, radionuclides such as 32P or 35S, or enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding HARC may be used for the diagnosis of conditions or diseases which are associated with expression of HARC. Examples of such conditions or diseases include cancer, autoimmune diseases, lymphoproliferative disorders, psoriasis, atherosclerosis, restenosis, AIDS, immunodeficiency diseases, ischemic injuries, neurodegenerative diseases, osteoporosis, myelodysplastic syndromes, toxin-induced diseases, cachexia, and viral infections. The polynucleotide sequences encoding HARC may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; or in dip stick, pin, ELISA or chip assays utilizing fluids or tissues from patient biopsies to detect altered HARC expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding HARC may be useful in assays that detect activation or induction of various cancers, particularly those mentioned above. The nucleotide sequences encoding HARC may be labeled by standard methods, and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the biopsied or extracted sample is significantly altered from that of a comparable control sample, the nucleotide sequences have hybridized with nucleotide sequences in the sample, and the presence of altered levels of nucleotide sequences encoding HARC in the sample indicates the presence of the associated disease. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or in monitoring the treatment of an individual patient.

In order to provide a basis for the diagnosis of disease associated with expression of HARC, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human,

with a sequence, or a fragment thereof, which encodes HARC, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with those from an experiment where a known amount of a substantially purified polynucleotide is used. Standard values obtained from normal samples may be compared with values obtained from samples from patients who are symptomatic for disease. Deviation between standard and subject values is used to establish the presence of disease.

Once disease is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to evaluate whether the level of expression in the patient begins to approximate that which is observed in the normal patient. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding HARC may involve the use of PCR. Such oligomers may be chemically synthesized, generated enzymatically, or produced from a recombinant source. Oligomers will preferably consist of two nucleotide sequences, one with sense orientation (5′→3′) and another with antisense (3′← 5′), employed under optimized conditions for identification of a specific gene or condition. The same two oligomers, nested sets of oligomers, or even a degenerate pool of oligomers may be employed under less stringent conditions for detection and/or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantitate the expression of HARC include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and standard curves onto which the experimental results are interpolated (Melby, P.C. et al. (1993) J. Immunol. Methods, 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem.

229-236). The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In another embodiment of the invention, the nucleic acid sequences which encode HARC may also be used to generate hybridization probes which are useful for mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome or to a specific region of the chromosome using well known techniques. Such techniques include FISH, FACS, or artificial chromosome constructions, such as yeast artificial chromosomes, bacterial artificial chromosomes, bacterial P1 constructions or single chromosome cDNA libraries as reviewed in Price, C.M. (1993) Blood Rev. 7:127-134, and Trask, B.J. (1991) Trends Genet. 7:149-154.

FISH (as described in Verma et al. (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York, NY) may be correlated with other physical chromosome mapping techniques and genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of the gene encoding HARC on a physical chromosomal map and a specific disease, or predisposition to a specific disease, may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier, or affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques such as linkage analysis using established chromosomal markers may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms, or parts thereof, by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, for example, AT to 11q22-23 (Gatti, R.A. et al. (1988) Nature 336:577-580), any sequences mapping to that area may represent associated or regulatory genes for further investigation. The nucleotide sequence of the subject invention

may also be used to detect differences in the chromosomal location due to translocation, inversion, etc. among normal, carrier, or affected individuals.

In another embodiment of the invention, HARC, its catalytic or immunogenic fragments or oligopeptides thereof, can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes, between HARC and the agent being tested, may be measured.

Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the protein of interest as described in published PCT application WO84/03564. In this method, as applied to HARC large numbers of different small test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The test compounds are reacted with HARC, or fragments thereof, and washed. Bound HARC is then detected by methods well known in the art. Purified HARC can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding HARC specifically compete with a test compound for binding HARC. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with HARC.

In additional embodiments, the nucleotide sequences which encode HARC may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

The examples below are provided to illustrate the subject invention and are not included for the purpose of limiting the invention.

INDUSTRIAL APPLICABILITY

I HUVENOB01 cDNA Library Construction

Incyte clone 036596 was identified in the HUVENOB01 cDNA library prepared from

control (or untreated) HUVEC cells. The HUVEC cell line is a normal, homogeneous, well-characterized early passage endothelial cell culture from human umbilical vein (Cell Systems Corp., Kirkland, WA).

The induced human endothelial cell (HUVEC) cDNA library was custom constructed by Stratagene (Cat. #937207). cDNA synthesis was primed using either oligo d(T) or random hexamers. Synthetic adaptor oligonucleotides were ligated onto cDNA ends enabling insertion into the lambda vector. The HUVEC cDNA library was constructed using the Uni-ZAP™ vector system (Stratagene). Finally, the two libraries were combined into a single library by mixing equal numbers of bacteriophage.

The custom-constructed library phage particles were infected into E. coli host strain XL1-Blue® (Stratagene) which has a high transformation efficiency, increasing the probability of obtaining rare, under-represented clones in the cDNA library. Alternative unidirectional vectors include, but are not limited to, pcDNAI (Invitrogen, San Diego, CA) and pSHlox-1 (Novagen, Madison, WI).

II Isolation and Sequencing of cDNA Clones

The phagemid forms of individual cDNA clones were obtained by the in vivo excision process in which the host bacterial strain was coinfected with both the lambda library phage and an f1 helper phage. Polypeptides derived from both the library-containing phage and the helper phage nicked the lambda DNA, initiated new DNA synthesis from defined sequences on the lambda target DNA and created a smaller, single-stranded circular phagemid DNA molecule that included all DNA sequences of the pBluescript® plasmid (Stratagene) and the cDNA insert. The phagemid DNA was secreted from the cells, purified, and used to re-infect fresh host cells where the double-stranded phagemid DNA was produced. Because the phagemid carries the gene for ß-lactamase, the newly-transformed bacteria were selected on media containing ampicillin.

Phagemid DNA was purified using the Magic Minipreps™ DNA Purification System (Cat. #A7100; Promega Corp., Madison, WI). This small-scale process provides a simple and reliable method for lysing the bacterial cells and rapidly isolating purified phagemid DNA using a proprietary DNA-binding resin.

Phagemid DNA was also purified using the QIAwell-8 Plasmid, QIAwell Plus and

QIAwell Ultra DNA purification systems (QIAGEN, Chatsworth, CA). This product line provides a convenient, rapid, and reliable high-throughput method for lysing the bacterial cells and isolating highly purified phagemid DNA. The DNA was eluted from the purification and binding resins already prepared for DNA sequencing and other analytical manipulations.

The cDNA inserts from random isolates of the induced and control HUVEC libraries were sequenced in part using the method of Sanger et al. (1975; J. Mol. Biol. 94:441f). Conventional enzymatic methods employ DNA polymerase Klenow fragments, Sequenase™ or Taq polymersae to extend DNA chains from an oligonucleotide primer annealed to the DNA template of interest. Methods have been developed for the use of both single- and double-stranded templates. The chain termination reaction products are separated by electrophoreses through urea-acrylamide gels and are detected by fluorescence.

III Homology Searching of cDNA Clones and Their Deduced Proteins

Each cDNA was compared to sequences in GenBank using two search algorithms. The first was originally developed by D.J. Lipman and W.R. Pearson. (1985; Science 227:1435-1441). In this algorithm, the homologous regions are searched in a two-step manner. In the first step, the highest homologous regions are determined by calculating a matching score using a homology score table. The parameter 'Ktup' is used in this step to establish the minimum window size to be shifted for comparing two sequences. Ktup also sets the number of bases that must match to extract the highest homologous region among the sequences. In this step, no insertion or deletions are applied, and the homology is displayed as an initial (INIT) value.

In the second step, the homologous regions are aligned to obtain the highest matching score by inserting a gap in order to add a probable deleted portion. The matching score obtained in the first step is recalculated using the homology score table and the insertion score table to an optimized (OPT) value in the final output.

DNA homologies between two sequences can be examined graphically using the Harr method of constructing dot matrix homology plots (Needleman, S.B. and C.O. Wunsch (1970) J. Mol. Biol. 48:443-453). This method produces a two-dimensional plot which can be useful in determining regions of homology versus regions of repetition.

The second algorithm was developed by Applied Biosystems and incorporated into the INHERIT™ 670 sequence analysis system. In this algorithm, Pattern Specification Language (TRW Inc., Los Angeles, CA) was used to determine regions of homology. The three parameters that determine how the sequence comparisons run were window size, window offset, and error tolerance. Using a combination of these three parameters, the DNA database was searched for sequences containing regions of homology to the query sequence, and the appropriate sequences were scored with an initial value. Subsequently, these homologous regions were examined using dot matrix homology plots to distinguish regions of homology from chance matches. Smith-Waterman alignments were used to display the results of the homology search.

Peptide and protein sequence homologies were ascertained using the INHERIT- 670 sequence analysis system using the methods similar to those used in DNA sequence homologies. Pattern Specification Language and parameter windows were used to search protein databases for sequences containing regions of homology which were scored with an initial value. Dot-matrix homology plots were examined to distinguish regions of significant homology from chance matches.

BLAST, which stands for Basic Local Alignment Search Tool (Altschul, S.F. (1993) J. Mol. Evol. 36:290-300; Altschul et al. (1990) J. Mol. Biol. 215:403-410), was used to search for local sequence alignments. BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs. BLAST is useful for matches which do not contain gaps. The fundamental unit of BLAST algorithm output is the High-scoring Segment Pair (HSP).

An HSP consists of two sequence fragments of arbitrary but equal lengths whose alignment is locally maximal and for which the alignment score meets or exceeds a threshold or cutoff score set by the user. The BLAST approach is to look for HSPs between a query sequence and a database sequence, to evaluate the statistical significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. The parameter E establishes the statistically significant threshold for reporting database sequence matches. E is interpreted as the upper bound of the expected frequency of

chance occurrence of an HSP (or set of HSPs) within the context of the entire database search. Any database sequence whose match satisfies E is reported in the program output.

IV Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound (Sambrook et al., supra).

Analogous computer techniques using BLAST (Altschul, S.F. 1993 and 1990, supra) are used to search for identical or related molecules in nucleotide databases such as GenBank or the LIFESEQ™ database (Incyte Pharmaceuticals). This analysis is much faster than multiple, membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or homologous.

The basis of the search is the product score which is defined as:

$$\frac{\% \text{ sequence identity} \times \% \text{ maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1-2% error; and at 70, the match will be exact. Homologous molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analysis are reported as a list of libraries in which the transcript encoding HARC occurs. Abundance and percent abundance are also reported. Abundance directly reflects the number of times a particular transcript is represented in a cDNA library, and percent abundance is abundance divided by the total number of sequences examined in the cDNA library.

V Extension of HARC-Encoding Polynucleotides to Full Length or to Recover Regulatory Sequences

Full length HARC-encoding nucleic acid sequence (SEQ ID NO:2) is used to design oligonucleotide primers for extending a partial nucleotide sequence to full length or for obtaining 5′ or 3′, intron or other control sequences from genomic libraries. One primer is synthesized to initiate extension in the antisense direction (XLR) and the other is synthesized

to extend sequence in the sense direction (XLF). Primers are used to facilitate the extension of the known sequence "outward" generating amplicons containing new, unknown nucleotide sequence for the region of interest. The initial primers are designed from the cDNA using OLIGO 4.06 (National Biosciences), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68°-72° C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations is avoided.

The original, selected cDNA libraries, or a human genomic library are used to extend the sequence; the latter is most useful to obtain 5′ upstream regions. If more extension is necessary or desired, additional sets of primers are designed to further extend the known region.

By following the instructions for the XL-PCR kit (Perkin Elmer) and thoroughly mixing the enzyme and reaction mix, high fidelity amplification is obtained. Beginning with 40 pmol of each primer and the recommended concentrations of all other components of the kit, PCR is performed using the Peltier Thermal Cycler (PTC200; M.J. Research, Watertown, MA) and the following parameters:

Step 1	94°C for 1 min (initial denaturation)
Step 2	65°C for 1 min
Step 3	68°C for 6 min
Step 4	94°C for 15 sec
Step 5	65°C for 1 min
Step 6	68°C for 7 min
Step 7	Repeat step 4-6 for 15 additional cycles
Step 8	94°C for 15 sec
Step 9	65°C for 1 min
Step 10	68°C for 7:15 min
Step 11	Repeat step 8-10 for 12 cycles
Step 12	72°C for 8 min
Step 13	4°C (and holding)

A 5-10 μl aliquot of the reaction mixture is analyzed by electrophoresis on a low concentration (about 0.6-0.8%) agarose mini-gel to determine which reactions were successful in extending the sequence. Bands thought to contain the largest products are selected and removed from the gel. Further purification involves using a commercial gel

extraction method such as QIAQuick™ (QIAGEN). After recovery of the DNA, Klenow enzyme is used to trim single-stranded, nucleotide overhangs creating blunt ends which facilitate religation and cloning.

After ethanol precipitation, the products are redissolved in 13 μl of ligation buffer, 1 μl T4-DNA ligase (15 units) and 1 μl T4 polynucleotide kinase are added, and the mixture is incubated at room temperature for 2-3 hours or overnight at 16°C. Competent E. coli cells (in 40 μl of appropriate media) are transformed with 3 μl of ligation mixture and cultured in 80 μl of SOC medium (Sambrook et al., supra). After incubation for one hour at 37°C, the whole transformation mixture is plated on Luria Bertani (LB)-agar (Sambrook et al., supra) containing 2x Carb. The following day, several colonies are randomly picked from each plate and cultured in 150 μl of liquid LB/2x Carb medium placed in an individual well of an appropriate, commercially-available, sterile 96-well microtiter plate. The following day, 5 μl of each overnight culture is transferred into a non-sterile 96-well plate and after dilution 1:10 with water, 5 μl of each sample is transferred into a PCR array.

For PCR amplification, 18 μl of concentrated PCR reaction mix (3.3x) containing 4 units of rTth DNA polymerase, a vector primer, and one or both of the gene specific primers used for the extension reaction are added to each well. Amplification is performed using the following conditions:

Step 1	94°C for 60 sec
Step 2	94°C for 20 sec
Step 3	55°C for 30 sec
Step 4	72°C for 90 sec
Step 5	Repeat steps 2-4 for an additional 29 cycles
Step 6	72°C for 180 sec
Step 7	4°C (and holding)

Aliquots of the PCR reactions are run on agarose gels together with molecular weight markers. The sizes of the PCR products are compared to the original partial cDNAs, and appropriate clones are selected, ligated into plasmid, and sequenced.

VI Labeling and Use of Hybridization Probes

Hybridization probes derived from SEQ ID NO:2 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base-pairs, is specifically described, essentially the same procedure is used with larger

cDNA fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 (National Biosciences), labeled by combining 50 pmol of each oligomer and 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham) and T4 polynucleotide kinase (DuPont NEN®, Boston, MA). The labeled oligonucleotides are substantially purified with Sephadex G-25 superfine resin column (Pharmacia & Upjohn). A portion containing 10^7 counts per minute of each of the sense and antisense oligonucleotides is used in a typical membrane based hybridization analysis of human genomic DNA digested with one of the following endonucleases (Ase I, Bgl II, Eco RI, Pst I, Xba 1, or Pvu II; DuPont NEN®).

The DNA from each digest is fractionated on a 0.7 percent agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham, NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. After XOMAT AR™ film (Kodak, Rochester, NY) is exposed to the blots in a Phosphoimager cassette (Molecular Dynamics, Sunnyvale, CA) for several hours, hybridization patterns are compared visually.

VII Antisense Molecules

Antisense molecules to the HARC-encoding sequence, or any part thereof, is used to inhibit in vivo or in vitro expression of naturally occurring HARC. Although use of antisense oligonucleotides, comprising about 20 base-pairs, is specifically described, essentially the same procedure is used with larger cDNA fragments. An oligonucleotide based on the coding sequences of HARC, as shown in Figures 1A and 1B, is used to inhibit expression of naturally occurring HARC. The complementary oligonucleotide is designed from the most unique 5′ sequence as shown in Figures 1A and 1B and used either to inhibit transcription by preventing promoter binding to the upstream nontranslated sequence or translation of an HARC-encoding transcript by preventing the ribosome from binding. Using an appropriate portion of the signal and 5′ sequence of SEQ ID NO:2, an effective antisense oligonucleotide includes any 15-20 nucleotides spanning the region which translates into the signal or 5′ coding sequence of the polypeptide as shown in Figures 1A and 1B.

VIII Expression of HARC

Expression of HARC is accomplished by subcloning the cDNAs into appropriate

vectors and transforming the vectors into host cells. In this case, the cloning vector, pSport, previously used for the generation of the cDNA library is used to express HARC in E. coli. Upstream of the cloning site, this vector contains a promoter for ß-galactosidase, followed by sequence containing the amino-terminal Met, and the subsequent seven residues of ß-galactosidase. Immediately following these eight residues is a bacteriophage promoter useful for transcription and a linker containing a number of unique restriction sites.

Induction of an isolated, transformed bacterial strain with IPTG using standard methods produces a fusion protein which consists of the first eight residues of ß-galactosidase, about 5 to 15 residues of linker, and the full length protein. The signal residues direct the secretion of HARC into the bacterial growth media which can be used directly in the following assay for activity.

IX Demonstration of HARC Activity

HARC can be expressed by transforming a mammalian cell line such as COS7, HeLa or CHO with an eukaryotic expression vector encoding HARC. Eukaryotic expression vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. The cells are incubated for 48-72 hours after transformation under conditions appropriate for the cell line to allow expression and accumulation of HARC.

Extracts containing solubilized proteins are prepared from cells expressing HARC by methods well known in the art. Portions of the extract containing HARC are applied to the wells of a polyacrylamide gel and electrophoresed at constant current using techniques well known in the art until a suitable tracking dye, such as bromophenol blue (Sigma), has migrated to the bottom of the gel. Appropriate control samples, prepared from extracts of untransformed cells and/or cells transformed with vector sequences alone, are electrophoresed in parallel lanes of the gel. Protein standards of known molecular weight (BioRad, Hercules, CA) are run in adjacent lanes to calibrate the gel.

The separated proteins are blotted onto a nitrocellulose membrane and assayed for binding to Ca^{2+} using the method described by Maruyama, K. et al. (1984; J. Biochem. (Tokyo) 95:511-519). After incubation in the presence of $^{45}Ca^{2+}$, the membrane is washed to remove nonspecifically bound radionuclide and exposed against Kodak XOMAT™ film for an appropriate period of time. A band will be visible on the film at a position that is

indicative of a protein of the size predicted for HARC. A band of similar mobility will not be present in control samples prepared from extracts of untransformed cells or cells transformed with vector sequence alone.

The presence of HARC in the band is confirmed using an antibody specific for HARC by the western blot procedure that is well known in the art. The membrane is incubated with HARC-specific antibodies derived from an animal such as rabbit. After washing to remove unbound antibodies, the membrane is incubated with tagged goat anti-rabbit immunoglobulins. The tag may consist of any of a number of chromogenic, fluorescent or enzymatic molecules that can be attached to immunoglobulins by techniques well known in the art. The presence of tagged goat immunoglobulins bound to rabbit anti-HARC antibodies may be detected using techniques appropriate to the nature of the tag. Pre-immune sera or unrelated antisera may be used as suitable controls for nonspecific binding to the membrane.

X Production of HARC Specific Antibodies

HARC that is substantially purified using PAGE electrophoresis (Sambrook, supra), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols. The amino acid sequence deduced from SEQ ID NO:2 is analyzed using DNASTAR software (DNASTAR Inc) to determine regions of high immunogenicity and a corresponding oligopolypeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions, is described by Ausubel et al. (supra), and others.

Typically, the oligopeptides are 15 residues in length, synthesized using an Applied Biosystems Peptide Synthesizer Model 431A using fmoc-chemistry, and coupled to keyhole limpet hemocyanin (KLH, Sigma, St. Louis, MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS; Ausubel et al., supra). Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. The resulting antisera are tested for antipeptide activity, for example, by binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radioiodinated, goat anti-rabbit IgG.

XI Purification of Naturally Occurring HARC Using Specific Antibodies

Naturally occurring or recombinant HARC is substantially purified by immunoaffinity chromatography using antibodies specific for HARC. An immunoaffinity column is constructed by covalently coupling HARC antibody to an activated chromatographic resin, such as CnBr-activated Sepharose (Pharmacia & Upjohn). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing HARC is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HARC (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/HARC binding (e.g., a buffer of pH 2-3 or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HARC is collected.

XII Identification of Molecules Which Interact with HARC

HARC or biologically active fragments thereof are labeled with ^{125}I Bolton-Hunter reagent (Bolton et al. (1973) Biochem. J. 133: 529). Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled HARC, washed and any wells with labeled HARC complex are assayed. Data obtained using different concentrations of HARC are used to calculate values for the number, affinity, and association of HARC with the candidate molecules.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: INCYTE PHARMACEUTICALS, INC.

(ii) TITLE OF THE INVENTION: HUMAN APOPTOSIS-RELATED CALCIUM-BINDING PROTEIN

(iii) NUMBER OF SEQUENCES: 3

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
(B) STREET: 3174 Porter Drive
(C) CITY: Palo Alto
(D) STATE: CA
(E) COUNTRY: US
(F) ZIP: 94304

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ Version 1.5

(vi) CURRENT APPLICATION DATA:
(A) PCT APPLICATION NUMBER: To Be Assigned
(B) FILING DATE: Filed Herewith

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/766,605
(B) FILING DATE: 12-DEC-1996

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Billings, Lucy J.
(B) REGISTRATION NUMBER: 36,749
(C) REFERENCE/DOCKET NUMBER: PF-0174 PCT

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 650-855-0555
(B) TELEFAX: 650-845-4166

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 214 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Met Phe Pro Xaa Gly Thr Pro Gly Xaa Pro Tyr Gly Xaa Ala Ala Pro
 1               5                  10                  15
Gly Gly Pro Tyr Gly Gln Pro Pro Pro Ser Ser Tyr Gly Xaa Gln Gln
            20                  25                  30
Pro Gly Xaa Tyr Gly Gln Gly Gly Ala Pro Pro Asn Val Asp Pro Glu
        35                  40                  45
Ala Tyr Ser Trp Phe Gln Ser Val Asp Ser Asp His Ser Gly Tyr Ile
    50                  55                  60
Ser Met Lys Glu Leu Lys Gln Ala Leu Val Asn Cys Asn Trp Ser Ser
65                  70                  75                  80
Phe Asn Asp Glu Thr Cys Leu Met Met Ile Asn Met Phe Asp Lys Thr
                85                  90                  95
Lys Ser Gly Arg Ile Asp Val Tyr Gly Phe Ser Ala Leu Trp Lys Phe
            100                 105                 110
Ile Gln Gln Trp Lys Asn Leu Phe Gln Gln Tyr Asp Arg Asp Arg Ser
        115                 120                 125
Gly Ser Ile Ser Tyr Thr Glu Leu Gln Gln Ala Leu Ser Gln Met Gly
    130                 135                 140
Tyr Asn Leu Ser Pro Gln Phe Thr Gln Leu Leu Val Ser Arg Tyr Cys
145                 150                 155                 160
Pro Arg Ser Ala Asn Pro Ala Met Gln Leu Asp Arg Phe Ile Gln Val
                165                 170                 175
Cys Thr Gln Leu Gln Val Leu Thr Glu Ala Phe Arg Glu Lys Asp Thr
            180                 185                 190
Ala Val Gln Gly Asn Ile Arg Leu Ser Phe Glu Asp Phe Val Thr Met
        195                 200                 205
Thr Ala Ser Arg Met Leu
    210
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 776 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:
(B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2.

```
TGGACACCCC AATCCTGGNA TGTTCCCCTN TGGAACTCCA GGAGNACCAT ATGGCGNTGC      60
AGCTCCCGGG GGCCCCTATG GTCAGCCACC TCCAAGTTCC TACGGTGNCC AGCAGCCTGG     120
GNTTTATGGA CAGGGTGGCG CCCCTCCCAA TGTGGATCCT GAGGCCTACT CCTGGTTCCA     180
GTCGGTGGAC TCAGATCACA GTGGCTATAT CTCCATGAAG GAGCTAAAGC AGGCCCTGGT     240
CAACTGCAAT TGGTCTTCAT TCAATGATGA GACCTGCCTC ATGATGATAA ACATGTTTGA     300
CAAGACCAAG TCAGGCCGCA TCGATGTCTA CGGCTTCTCA GCCCTGTGGA AATTCATCCA     360
GCAGTGGAAG AACCTCTTCC AGCAGTATGA CCGGGACCGC TCGGGCTCCA TTAGCTACAC     420
AGAGCTGCAG CAAGCTCTGT CCCAAATGGG CTACAACCTG AGCCCCCAGT TCACCCAGCT     480
TCTGGTCTCC CGCTACTGCC CACGCTCTGC CAATCCTGCC ATGCAGCTTC NCCGCTTCAT     540
CCAGGTGTGC ACCCAGCTGC AGGTGCTGAC AGAGGCCTTC CGGGAGAAG  CACAGCTGT     600
ACAAGGCAAC ATCCGGCTCA GCTTCGAGGA CTTCGTCACC ATGACAGCT  TCGGATGCT      660
ATGACCCAAC CATCTGTGGA GAGTGGAGTG CACCAGGGAC CTTTCCTGGN TCTTAGAGT      720
GAGAGAAGTA TGTGGACATC TCTTCTTTTC CTGTCCCTCT AGAAGAACAT TCTCCC         776
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 191 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: GenBank
(B) CLONE: 1213520

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Met Ala Ala Tyr Ser Tyr Arg Pro Gly Pro Gly Gly Gly Pro Gly Pro
 1               5                  10                  15
Ala Ala Gly Ala Ala Leu Pro Asp Gln Ser Phe Leu Trp Asn Val Phe
            20                  25                  30
Gln Arg Val Asp Lys Asp Arg Ser Gly Val Ile Ser Asp Asn Glu Leu
        35                  40                  45
Gln Gln Ala Leu Ser Asn Gly Thr Trp Thr Pro Phe Asn Pro Val Thr
    50                  55                  60
Val Arg Ser Ile Ile Ser Met Phe Asp Arg Glu Asn Lys Ala Gly Val
65                  70                  75                  80
Asn Phe Ser Glu Phe Thr Gly Val Trp Lys Tyr Ile Thr Asp Trp Gln
                85                  90                  95
Asn Val Phe Arg Thr Tyr Asp Arg Asp Asn Ser Gly Met Ile Asp Lys
            100                 105                 110
Asn Glu Leu Lys Gln Ala Leu Ser Gly Phe Gly Tyr Arg Leu Ser Asp
        115                 120                 125
Gln Phe His Asp Ile Leu Ile Arg Lys Phe Asp Arg Gln Gly Arg Gly
    130                 135                 140
Gln Ile Ala Phe Asp Asp Phe Ile Gln Gly Cys Ile Val Leu Gln Arg
145                 150                 155                 160
Leu Thr Asp Ile Phe Arg Arg Tyr Asp Thr Asp Gln Asp Gly Trp Ile
                165                 170                 175
Gln Val Ser Tyr Glu Gln Tyr Leu Ser Met Val Phe Ser Ile Val
            180                 185                 190
```

What is claimed is:

1. A substantially purified human apoptosis-related calcium-binding protein (HARC) comprising the amino acid sequence of SEQ ID NO:1 or fragments thereof.

2. An isolated and purified polynucleotide sequence encoding the human apoptosis-related calcium-binding protein of claim 1.

3. A polynucleotide sequence which hybridizes under stringent conditions to the polynucleotide sequence of claim 2.

4. A hybridization probe comprising the polynucleotide sequence of claim 2.

5. An isolated and purified polynucleotide sequence comprising SEQ ID NO:2 or variants thereof.

6. A polynucleotide sequence which is complementary to SEQ ID NO:2 or variants thereof.

7. A hybridization probe comprising the polynucleotide sequence of claim 6.

8. An expression vector containing the polynucleotide sequence of claim 2.

9. A host cell containing the vector of claim 8.

10. A method for producing a polypeptide comprising the amino acid sequence of SEQ ID NO:1, the method comprising the steps of:

a) culturing the host cell of claim 9 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

11. A pharmaceutical composition comprising a substantially purified human apoptosis-related calcium-binding protein having an amino acid sequence of SEQ ID NO:1 in conjunction with a suitable pharmaceutical carrier.

12. A purified antibody which binds specifically to the polypeptide of claim 1.

13. A purified agonist which specifically binds to and modulates the activity of the polypeptide of claim 1.

14. A method for treating disease resulting from decreased apoptosis comprising administering to a subject in need of such treatment an effective amount of the agonist of claim 13.

15. A purified antagonist which specifically binds to and modulates the activity of the polypeptide of claim 1.

16. A method for treating disease resulting from increased apoptosis comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 15.

17. A method for detecting polynucleotides encoding the human apoptosis-related calcium-binding protein in a biological sample comprising the steps of:

a) hybridizing a polynucleotide of claim 6 to nucleic acid material of a biological sample, thereby forming a hybridization complex; and

b) detecting said hybridization complex, wherein the presence of said complex correlates with the presence of a polynucleotide encoding human apoptosis-related calcium-binding protein in said biological sample.

18. The method of claim 17 wherein before hybridization, the nucleic acid material of the biological sample is amplified by the polymerase chain reaction.

```
                9          18          27          36          45          54
5'   T GGA CAC CCC AAT CCT GGN ATG TTC CCC TNT GGA ACT CCA GGA GNA CCA TAT
                               M   F   P   X   G   T   P   G   X   P   Y

               63          72          81          90          99         108
   GGC GNT GCA GCT CCC GGG GGC CCC TAT GGT CAG CCA CCT CCA AGT TCC TAC GGT
   G   Y   A   A   P   G   G   P   Y   G   Q   P   P   P   S   S   Y   G

              117         126         135         144         153         162
   GNC CAG CAG CCT GGG NTT TAT GGA CAG GGT GGC GCC CCT CCC AAT GTG GAT CCT
   X   Q   Q   P   G   X   Y   G   Q   G   G   A   P   P   N   V   D   P

              171         180         189         198         207         216
   GAG GCC TAC TCC TGG TTC CAG TCG GTG GAC TCA GAT CAC AGT GGC TAT ATC TCC
   E   A   Y   S   W   F   Q   S   V   D   S   D   H   S   G   Y   I   S

              225         234         243         252         261         270
   ATG AAG GAG CTA AAG CAG GCC CTG GTC AAC TGC AAT TGG TCT TCA TTC AAT GAT
   M   K   E   L   K   Q   A   L   V   N   C   N   W   S   S   F   N   D

              279         288         297         306         315         324
   GAG ACC TGC CTC ATG ATG ATA AAC ATG TTT GAC AAG ACC AAG TCA GGC CGC ATC
   E   T   C   L   M   M   I   N   M   F   D   K   T   K   S   G   R   I

              333         342         351         360         369         378
   GAT GTC TAC GGC TTC TCA GCC CTG TGG AAA TTC ATC CAG CAG TGG AAG AAC CTC
   D   V   Y   G   F   S   A   L   W   K   F   I   Q   Q   W   K   N   L
```

FIGURE 1A

```
          387        396        405        414        423        432
TTC CAG CAG TAT GAC CGG GAC CGC TCG GGC TCC ATT AGC TAC ACA GAG CTG CAG
F   Q   Q   Y   D   R   D   R   S   G   S   I   S   Y   T   E   L   Q

          441        450        459        468        477        486
CAA GCT CTG TCC CAA ATG GGC TAC AAC CTG AGC CCC CAG TTC ACC CAG CTT CTG
Q   A   L   S   Q   M   G   Y   N   L   S   P   Q   F   T   Q   L   L

          495        504        513        522        531        540
GTC TCC CGC TAC TGC CCA CGC TCT GCC AAT CCT GCC ATG CAG CTT GAC CGC TTC
V   S   R   Y   C   P   R   S   A   N   P   A   M   Q   L   D   R   F

          549        558        567        576        585        594
ATC CAG GTG TGC ACC CAG CTG CAG GTG CTG ACA GAG GCC TTC CGG GAG AAG GAC
I   Q   V   C   T   Q   L   Q   V   L   T   E   A   F   R   E   K   D

          603        612        621        630        639        648
ACA GCT GTA CAA GGC AAC ATC CGG CTC AGC TTC GAG GAC TTC GTC ACC ATG ACA
T   A   V   Q   G   N   I   R   L   S   F   E   D   F   V   T   M   T

          657        666        675        684        693        702
GCT TCT CGG ATG CTA TGA CCC AAC CAT CTG TGG AGA GTG GAG TGC ACC AGG GAC
A   S   R   M   L   *
          711        720        729        738        747        756
CTT TCC TGG CTT CTT AGA GTG AGA GAA GTA TGT GGA CAT CTC TTC TTT TCC TGT
          765        774
CCC TCT AGA AGA ACA TTC TCC C 3'
P   S   R   R   T   F   S
```

FIGURE 1B

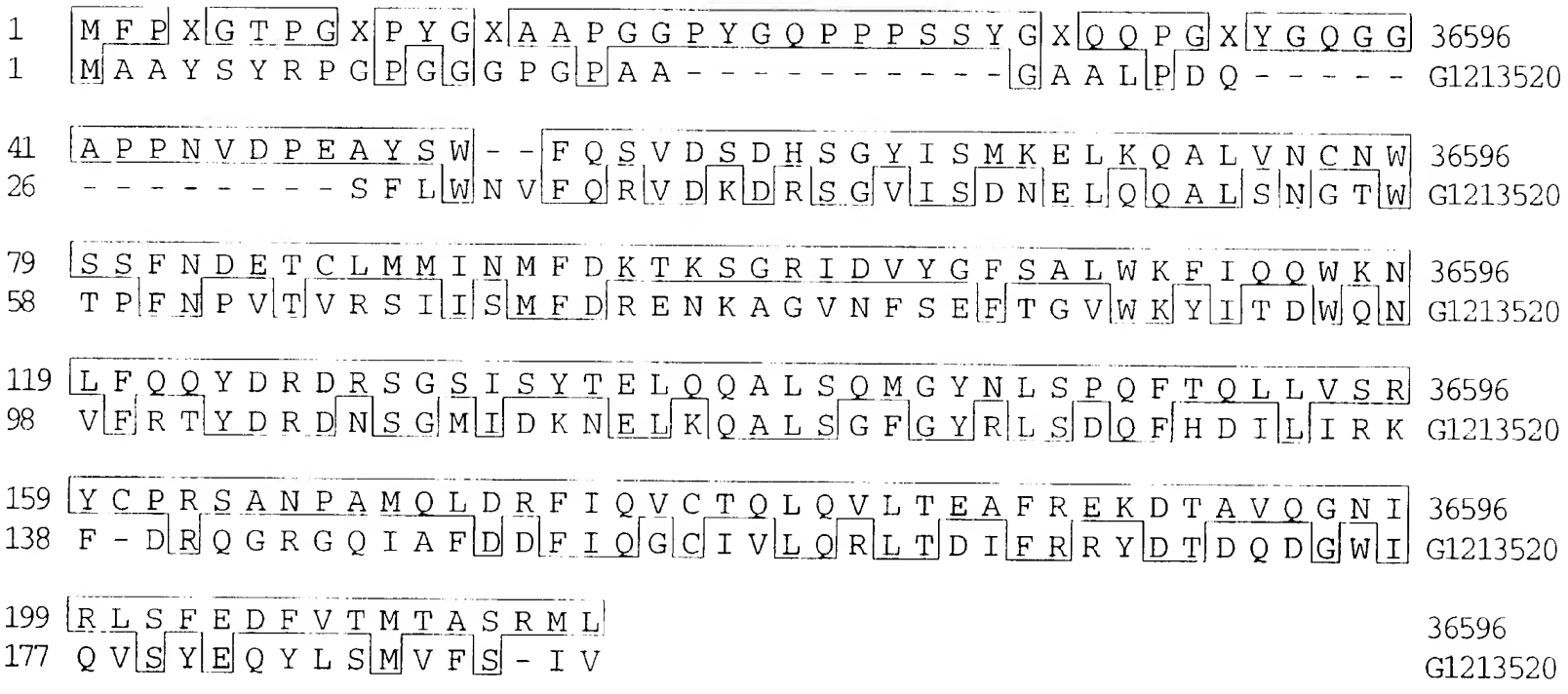


FIGURE 2

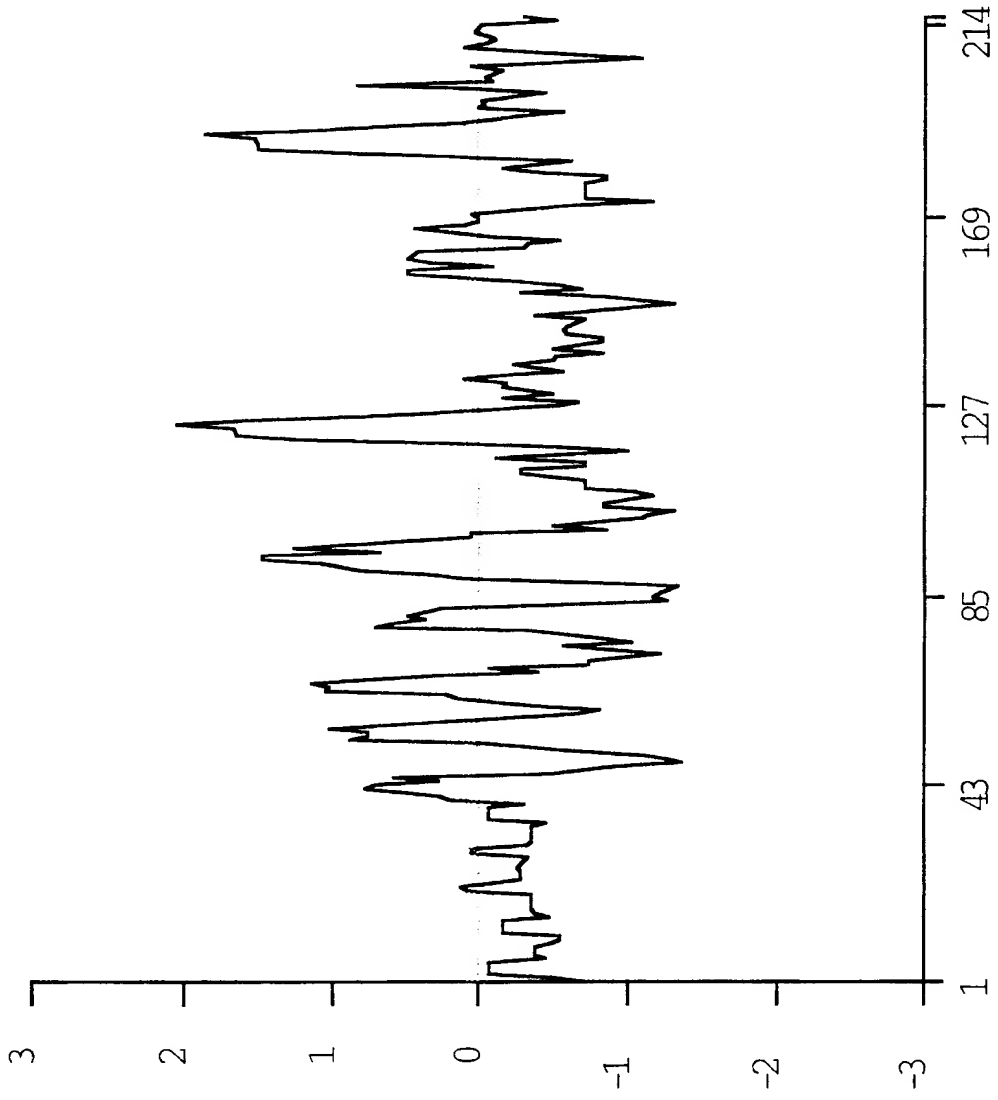


FIGURE 3

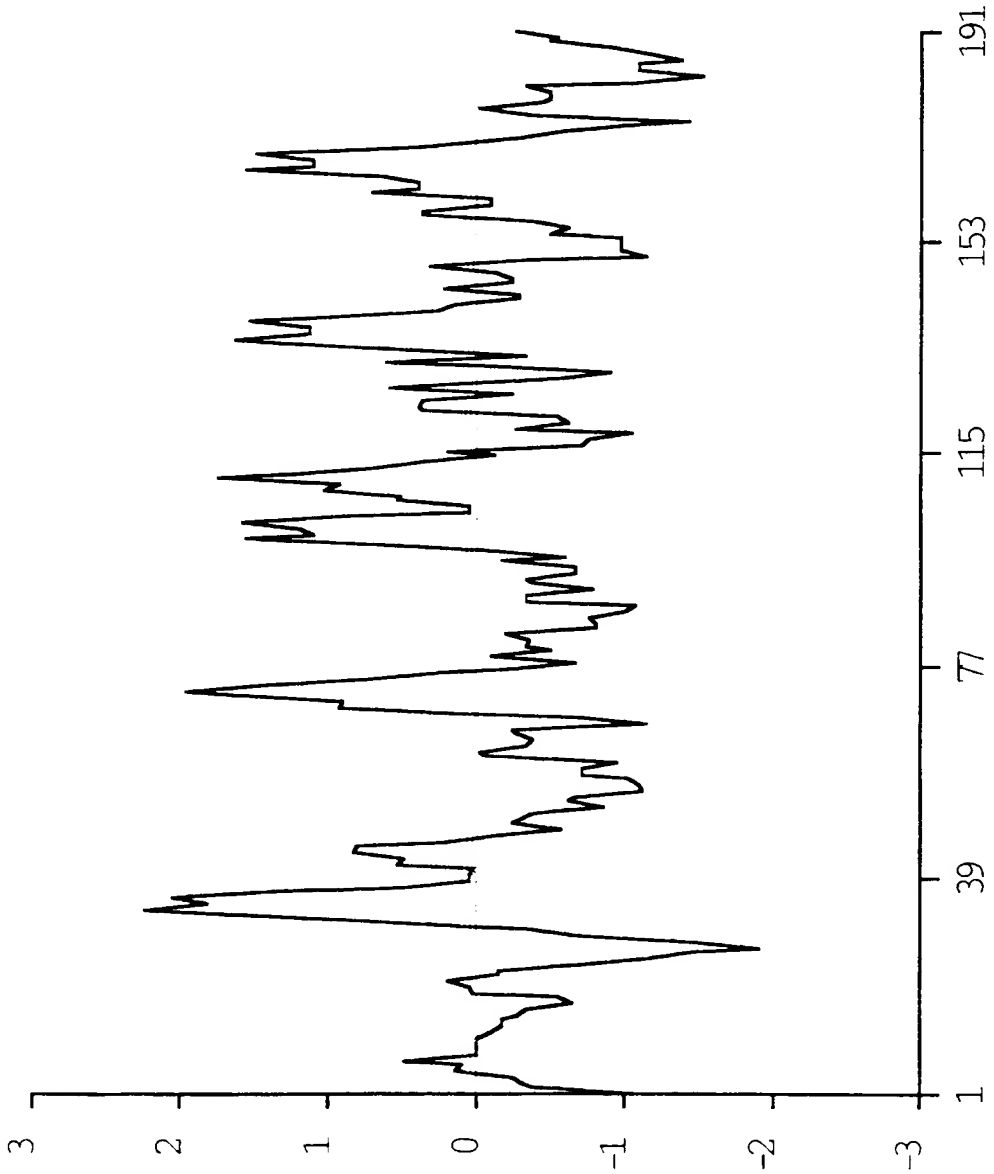


FIGURE 4

Library	Lib Description	Abun	Pct Abun
TESTNOT04	testis, 37 M	1	0.1075
HNT3AZT01	hNT2 cell line, teratocarcinoma, treated AZ	1	0.0685
ADRENOT03	adrenal gland, 17 M	2	0.0682
HIPONON01	brain, hippocampus, 72 F, NORM	2	0.0675
ENDCNOT02	endothelial cells, dermal microvascular, 30 F	1	0.0604
LIVRTUT01	liver tumor, metastasis, 51 F	2	0.0518
PROSNOT18	prostate, 58 M	2	0.0513
MUSCNOT01	muscle, skeletal	1	0.0445
COLNNOT08	colon, 60 M	1	0.0426
HUVENOB01	HUVEC endothelial cell line, control	1	0.0420
BRAINON01	brain, 26 M, NORM	1	0.0417
CRBLNOT01	brain, cerebellum, 69 M	2	0.0391
TONSNOT01	tonsil, hyperplasia, 6 M	1	0.0339
PLACNOT02	placenta, fetal F	2	0.0336
PROSNOT20	prostate, 65 M, match to PROSTUT12	1	0.0336
BSTMNON02	brain stem, 72 M, NORM	1	0.0319
LUNGTUT03	lung tumor, 69 M, match to LUNGNOT15	2	0.0318
THYRTUT03	thyroid tumor, benign, 17 M	1	0.0276
BRSTNOM01	breast, F, NORM, WM	1	0.0264
STOMFET01	stomach, fetal F	1	0.0255
HIPONOT01	brain, hippocampus, 72 F	1	0.0239
THYRNOT01	thyroid, 64 F	1	0.0229
CARDFEM01	heart, fetal, NORM, WM	2	0.0223
PANCTUT02	pancreatic tumor, carcinoma, 45 F	1	0.0202

FIGURE 5A

BRSTTUT01	breast tumor, 55 F, match to BRSTNOT02	2	0.0189
KERANOT02	keratinocytes, primary cell line, 30 F	1	0.0182
SINTBST01	small intestine, ileum, Crohn's, 18 F	1	0.0168
PLACNOM02	placenta, neonatal F, NORM, WM	3	0.0167
UTRSNOT02	uterus, 34 F	1	0.0167
LUNGFEM01	lung, fetal, NORM, WM	1	0.0148
BRSTNOT03	breast, 54 F, match to BRSTTUT02	1	0.0147
COLNFET02	colon, fetal F	1	0.0143
THYRNOT03	thyroid tumor, adenoma, 28 F	1	0.0138
BRAITUT01	brain tumor, oligoastrocytoma, 50 F	1	0.0134
TESTNOT03	testis, 37 M	1	0.0129
BLADTUT04	bladder tumor, 60 M, match to BLADNOT05	1	0.0127
SPLNFET02	spleen, fetal M	1	0.0126
PROSTUT04	prostate tumor, 57 M, match to PROSNOT06	1	0.0117
PROSNOT06	prostate, 57 M, match to PROSTUT04	1	0.0114
KIDNNOT05	kidney, neonatal F	1	0.0106
EOSIHET02	eosinophils, hypereosinophilia, 48 M	1	0.0105
LUNGAST01	lung, asthma, 17 M	1	0.0094
UCMCL5T01	mononuclear cells, treated IL-5	1	0.0084
LIVSFEM02	liver/spleen, fetal M, NORM, WM	3	0.0079

FIGURE 5B

INTERNATIONAL SEARCH REPORT

[redacted]nal Application No
PCT/US 97/22541

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/47 C07K16/18 A61K38/17 C12Q1/68
G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EMBL Databank Accession number h72140 Hillier L. et al. 02-11-1995 XP002062051 see the whole document ---	1-10
X	EMBL Databank Accession number H91107 Hillier L. et al. 30-11-95 XP002062052 see the whole document ---	1-10

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

3

Date of the actual completion of the international search	Date of mailing of the international search report
14 April 1998	15. 05. 98

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Gurdjian, D

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
Y	MAKI M (REPRINT) ET AL: "A growing family of the Ca2+-binding proteins with five EF-hand motifs" BIOCHEMICAL JOURNAL, 1997, 328, 718-720, XP002062049 see the whole document ---	1-13,15, 17
Y	VITO P ET AL: "Interfering with apoptosis: Ca(2+)-binding protein ALG-2 and Alzheimer's disease gene ALG-3." SCIENCE, JAN 26 1996, 271 (5248) P521-5, UNITED STATES, XP002062050 cited in the application see the whole document ---	1-13,15, 17
Y	WO 94 21817 A (US HEALTH) 29 September 1994 see claims 10-24 -----	11-13, 15,17

3

INTERNATIONAL SEARCH REPORT

rnational application No
PCT/US 97/22541

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

see FURTHER INFORMATION sheet PCT/ISA/210

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Remark : Although claims 14 16 are directed to a method of treatment of the human/animal body , the search has been carried out and based on the alleged effects of the compound/composition.

INTERNATIONAL SEARCH REPORT

[redacted]ion on patent family members

[redacted]ial Application No
PCT/US 97/22541

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9421817 A	29-09-94	US 5607831 A	04-03-97
		AU 6367494 A	11-10-94

CORRECTED VERSION*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/47, 16/18, A61K 38/17, C12Q 1/68, G01N 33/68

A1

(11) International Publication Number: WO 98/26068

(43) International Publication Date: 18 June 1998 (18.06.98)

(21) International Application Number: PCT/US97/22541

(22) International Filing Date: 8 December 1997 (08.12.97)

(30) Priority Data:
08/766,605 12 December 1996 (12.12.96) US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US 08/766,605 (CIP)
Filed on 12 December 1996 (12.12.96)

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #12, Mountain View, CA 94040 (US). GOLI, Surya, K. [IN/US]; 620 Iris Avenue #338, Sunnyvale, CA 94086 (US).

(74) Agent: BILLINGS, Lucy, J.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AT, AU, BR, CA, CH, CN, DE, DK, ES, FI, GB, IL, JP, KR, MX, NO, NZ, RU, SE, SG, US, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: HUMAN APOPTOSIS-RELATED CALCIUM-BINDING PROTEIN

```
1   MFPXGTPGXPYGXAAPGGPYGQPPPSSYGXQQPGXYGQGG  36596
1   MAAYSYRPGPGGGPGPAA----------GAALPDQ-----  G1213520

41  APPNVDPEAYSW--FQSVDSDHSGYISMKELKQALVNCNW  36596
26  ---------SFLWNVFQRVDKDRSGVISDNELQQALSNGTW  G1213520

79  SSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKN  36596
58  TPFNPVTVRSIISMFDRENKAGVNFSEFTGVWKYITDWQN  G1213520

119 LFQQYDRDRSGSISYTELQQALSQMGYNLSPQFTQLLVSR  36596
98  VFRTYDRDNSGMIDKNELKQALSGFGYRLSDQFHDILIRK  G1213520

159 YCPRSANPAMQLDRFIQVCTQLQVLTEAFREKDTAVQGNI  36596
138 F-DRQGRGQIAFDDFIQGCIVLQRLTDIFRRYDTDQDGWI  G1213520

199 RLSFEDFVTMTASRML                          36596
177 QVSYEQYLSMVFS-IV                          G1213520
```

(57) Abstract

The present invention provides a human apoptosis-related calcium-binding protein (HARC) and polynucleotides which identify and encode HARC. The invention also provides genetically engineered expression vectors and host cells comprising the nucleic acid sequences encoding HARC and a method for producing HARC. The invention also provides for agonists, antibodies, or antagonists specifically binding HARC, and their use, in the prevention and treatment of diseases associated with expression of HARC. Additionally, the invention provides for the use of antisense molecules to polynucleotides encoding HARC for the treatment of diseases associated with the expression of HARC. The invention also provides diagnostic assays which utilize the polynucleotide, or fragments or the complement thereof, and antibodies specifically binding HARC.

*(Referred to in PCT Gazette No. 38/1998, Section II)

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		